

The non-linear infrared-radio correlation of low-z galaxies: implications for redshift evolution, a new radio SFR recipe, and how to minimize selection bias

Article (Accepted Version)

Molnar, Daniel Cs, Sargent, Mark T, Leslie, Sarah, Magnelli, Benjamin, Schinnerer, Eva, Zamorani, Giovanni, Delhaize, Jacinta, Smolcic, Vernesa, Tisanic, Kresimir and Vardoulaki, Eleni (2021) The non-linear infrared-radio correlation of low-z galaxies: implications for redshift evolution, a new radio SFR recipe, and how to minimize selection bias. Monthly Notices of the Royal Astronomical Society. ISSN 0035-8711

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/97999/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

The non-linear infrared-radio correlation of low-*z* galaxies: implications for redshift evolution, a new radio SFR recipe, and how to minimize selection bias

Dániel Cs. Molnár,^{1,2,3*} Mark T. Sargent,¹ Sarah Leslie,^{4,3} Benjamin Magnelli,⁵ Eva Schinnerer,³ Giovanni Zamorani,⁶ Jacinta Delhaize,⁷ Vernesa Smolčić,⁸ Krešimir Tisanić,⁸ Eleni Vardoulaki^{5,9,10}

¹*Astronomy Centre, Department of Physics & Astronomy, University of Sussex, Brighton, BN1 9QH, England*

²*INAF - Osservatorio Astronomico di Cagliari, Via della Scienza 5, I-09047 Selargius (CA), Italy*

³*MPI for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*

⁴*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

⁵*Argelander Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, Bonn, D-53121, Germany*

⁶*INAF-Osservatorio Astronomico di Bologna, via P. Gobetti 93/3, 40129 Bologna, Italy*

⁷*Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa*

⁸*Department of Physics, Faculty of Science, University of Zagreb, Bijenička cesta 32, 10000 Zagreb, Croatia*

⁹*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany*

¹⁰*Thüringer Landessternwarte, Sternwarte 5, 07778 Tautenburg, Germany*

Accepted 2021 March 05. Received 2021 March 05; in original form 2019 December 07

ABSTRACT

The infrared-radio correlation (IRRC) underpins many commonly used radio luminosity–star formation rate (SFR) calibrations. In preparation for the new generation of radio surveys we revisit the IRRC of low-*z* galaxies by (a) drawing on the best currently available IR and 1.4 GHz radio photometry, plus ancillary data over the widest possible area, and (b) carefully assessing potential systematics. We compile a catalogue of $\sim 9,500$ $z < 0.2$ galaxies and derive their 1.4 GHz radio ($L_{1.4}$), total IR, and monochromatic IR luminosities in up to seven bands, allowing us to parameterize the wavelength-dependence of monochromatic IRRCs from 22–500 μm . For the first time for low-*z* samples, we quantify how poorly matched IR and radio survey depths bias measured median IR/radio ratios, \bar{q}_{TIR} , and discuss the level of biasing expected for low-*z* IRRC studies in ASKAP/MeerKAT fields. For our subset of $\sim 2,000$ high-confidence star-forming galaxies we find a median \bar{q}_{TIR} of 2.54 (scatter: 0.17 dex). We show that \bar{q}_{TIR} correlates with $L_{1.4}$, implying a non-linear IRRC with slope 1.11 ± 0.01 . Our new $L_{1.4}$ –SFR calibration, which incorporates this non-linearity, reproduces SFRs from panchromatic SED fits substantially better than previous IRRC-based recipes. Finally, we match the evolutionary slope of recently measured \bar{q}_{TIR} –redshift trends without having to invoke redshift evolution of the IRRC. In this framework, the redshift evolution of \bar{q}_{TIR} reported at GHz frequencies in the literature is the consequence of a partial, redshift-dependent sampling of a non-linear IRRC obeyed by low-*z* and distant galaxies.

Key words: galaxies: star formation – radio continuum: galaxies – infrared: galaxies

1 INTRODUCTION

The infrared and radio synchrotron continuum luminosities are observed to be closely related in star-forming galaxies (van der Kruit 1971, 1973; de Jong et al. 1985; Helou et al.

1985; Condon 1992; Yun et al. 2001). Since the far-infrared (FIR; 25 – 1000 μm) emission is predominantly generated by star formation (SF) activity (Kennicutt 1998; Charlot & Fall 2000), this so-called infrared-radio correlation (IRRC) implies that radio power in most galaxies is also related to SF. The IRRC has been used to establish a radio-based star formation rate (SFR) calibration (e.g. Condon 1992; Murphy

* E-mail: daniel.molnar@inaf.it

et al. 2011). The main advantages of the radio synchrotron continuum over other SF tracers are (i) the fact that it is unattenuated by interstellar dust, and hence does not require appropriate corrections, (ii) the high angular resolution that is achievable in interferometric observations with radio telescope arrays, and (iii) especially with next generation telescopes, superb sensitivity and survey speed. However, despite abundant literature on the topic (e.g. Voelk 1989; Helou & Bicay 1993; Bell 2003; Lacki et al. 2010; Schleicher & Beck 2013), the detailed physics shaping the IRRC remain poorly understood from the theoretical perspective.

In order to better leverage the aforementioned strengths of radio continuum emission as an SF tracer, numerous studies in the past decade have sought to improve its accuracy by calibrating it against other, theoretically better-established SF tracers (e.g. Hodge et al. 2008; Brown et al. 2017; Davies et al. 2017; Gürkan et al. 2018; Read et al. 2018; Duncan et al. 2020), or examined the variation of the IRRC with other galaxy properties, such as stellar mass (e.g. Magnelli et al. 2015; Delvecchio et al. 2020) or galaxy type (Morić et al. 2010; Roychowdhury & Chengalur 2012; Nyland et al. 2017). An especially frequently debated aspect of the IRRC is its (non-)evolution with redshift (e.g. Garrett 2002; Appleton et al. 2004; Garn et al. 2009; Jarvis et al. 2010; Sargent et al. 2010a,b; Mao et al. 2011; Smith et al. 2014; Magnelli et al. 2015; Calistro Rivera et al. 2017; Delhaize et al. 2017; Molnár et al. 2018; Delvecchio et al. 2020). A majority of these studies compare their results to the classical works of Yun et al. (2001) and Bell (2003), since these are considered to be the main reference points for the low- z IRRC. The overall IRRC properties, such as the slope and dispersion of the relation, proved to be broadly consistent between these two studies. With evidence for non-linearity at low IR luminosities supported by the findings of Yun et al. (2001), Bell (2003) provided a refined luminosity-dependent radio – SFR calibration. However, both of these cornerstone papers use the 60 and 100 μm photometry from the Infrared Astronomical Satellite (IRAS; Neugebauer et al. 1984) to estimate IR luminosities, and thus lack the now standard spectral energy distribution (SED) fitting approach, and had to rely on the shallow but wide radio coverage of the NRAO VLA Sky Survey (NVSS; Condon et al. 1998).

Since the publication of the seminal Yun et al. (2001) and Bell (2003) works, deeper radio and IR measurements and better overall IR photometric coverage have become available, mainly thanks to the Faint Images of the Radio Sky at Twenty centimetres survey (FIRST; Becker et al. 1995; Helfand et al. 2015), the Herschel Space Observatory (Pilbratt et al. 2010 hereafter, Herschel) and the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010). Furthermore, as discussed in Sargent et al. (2010a), IR- and radio-selection effects can bias median IR-radio ratio measurements. Avoiding such biases requires a careful approach to sample selection, and this has so far almost exclusively been discussed in the context of redshift evolution, but much less when it comes to calibrating radio-based SFR measurements on low-redshift samples. Meanwhile, a new generation of deeper and wider surveys on modern radio telescopes – e.g., SKA pathfinders, the Karl G. Jansky Very Large Array (JVLA; Perley et al. 2011) and the Low Frequency Array (LOFAR; van Haarlem et al. 2013), and SKA precursors, the Australian Square Kilometre Array Pathfinder

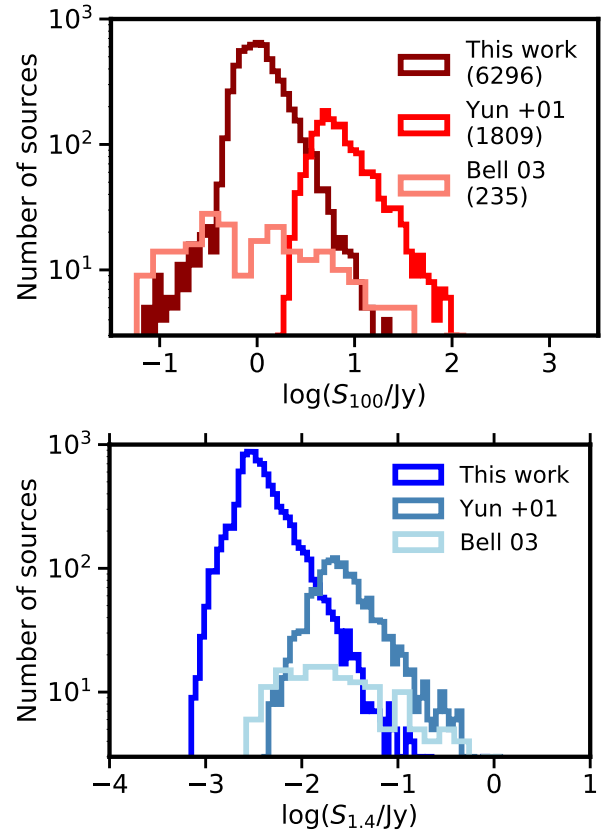


Figure 1. Comparison of 100 μm and 1.4 GHz radio flux densities in our catalogue to the flux distributions of the Yun et al. (2001) and Bell (2003) samples. Number counts are on a logarithmic scale to ease comparison between the samples.

(ASKAP; Johnston et al. 2007; DeBoer et al. 2009) and the Meer Karoo Array Telescope (MeerKAT; Booth et al. 2009) – is providing a more complete census of radio emission from star-forming galaxies both in the local and distant Universe. In preparation for this next generation of studies, it is thus timely to revisit the low- z IRRC. To this end we use the aforementioned FIRST, Herschel and WISE observations and other ancillary data. We also define a highly pure star-forming galaxy (SFG) sample through careful separation of SFGs and active galactic nuclei (AGN), and we perform SED fits that exploit IR photometry covering a broader wavelength range, to assemble a large $z < 0.2$ sample with an eye to quantifying potential systematics due to methodology and/or selection effects.

This work presents an initially ~ 5 times larger and ~ 4 times deeper data set than the one used in Yun et al. (2001). Fig. 1 shows a comparison of IR and radio flux densities in the catalogues of Bell (2003)¹, Yun et al. (2001) and our work. We probe fainter sources than Yun et al. (2001) both in the radio and IR, while we have a similar coverage to Bell (2003) at 100 μm . However, comparisons of both 1.4 GHz radio continuum ($L_{1.4\text{GHz}}$) and total IR (L_{TIR} ; 8 – 1000

¹ The 100 μm flux densities from Bell (2003) were derived using the published 60 μm IR luminosities and IRAS 60 and 100 μm flux density ratios.

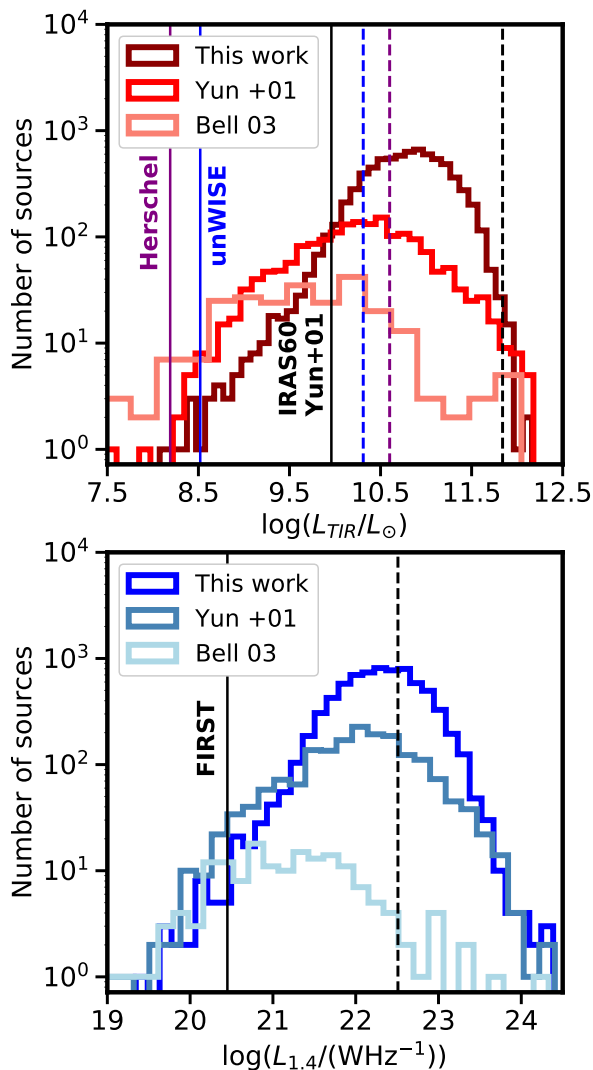


Figure 2. Total infrared and 1.4 GHz radio luminosity distributions of the Yun et al. (2001) and Bell (2003) samples in comparison to our catalogue (for details on the calculations of radio and IR luminosities, see Sect. 3.1.1). Solid (dashed) lines show the various luminosity limits at $z = 0.01$ ($z = 0.1$) derived from Fig. 7. Number counts are on a logarithmic scale to ease comparison between the samples.

μm) luminosities², seen in Fig. 2, reveal that we only substantially increase the number of high-luminosity objects at $\log(L_{1.4\text{GHz}}/\text{WHz}^{-1}) \geq 21$ and $\log(L_{\text{TIR}}/L_{\odot}) \geq 10$, respectively. The primary reason for this is the ~ 2.6 times larger area covered by Yun et al. (2001). Our catalogue with IR and radio luminosity measurements for 9,645 galaxies is publicly available (for details see Appendix C) to support follow-up

² We converted the publicly available 1.4 GHz luminosities of Yun et al. (2001) and Bell (2003) to flux densities assuming a radio spectral slope of -0.7 . Total IR luminosities for the sample in Yun et al. (2001) were calculated by first using the published 60 and $100\mu\text{m}$ flux densities and Eq. (2) and (3) in Yun et al. (2001) to obtain FIR luminosities, and then multiplying them by ~ 2 , the average offset between FIR and TIR luminosities (see e.g. Bell 2003).

studies investigating the low- z IRRC’s dependence on various galaxy parameters and thus gaining insights into the physics regulating the correlation.

In Sect. 2 we describe the archival data products used, and the catalogue construction process. Sect. 3 details the calculation of IR and radio luminosities, the identification of AGN and SFG sources, and gives a brief summary of all the data products used in our analysis. In Sect. 4 we characterise the properties of both the monochromatic and bolometric IRRCs of low- z galaxies, and we demonstrate and quantify sensitivity related selection effects. Based on this, in Sect. 5, we discuss the implications of our findings for the radio–SFR calibration, and for interpreting the observed redshift evolution of the IRRC.

Throughout this paper, we use a flat ΛCDM cosmology with $\Omega_M = 0.3$ and $H_0 = 70\text{ km s}^{-1}\text{ Mpc}^{-1}$. Star formation rates reported assume a Chabrier (2003) initial mass function³ (IMF).

2 DATA

The starting point for our sample construction is the Sloan Digital Sky Survey (SDSS) Data Release 12 (Alam et al. 2015), which provides positions and redshifts of nearly 470 million unique optical sources over roughly 1/3 of the Celestial sphere. For our analysis, we selected SDSS galaxies with spectroscopic (58%) or, where unavailable, photometric (42%) redshifts⁴ below $z = 0.2$, corresponding to a look-back time of $\lesssim 2.4$ Gyr. This selection results in an optical parent sample of 3,001,410 galaxies for which we identify infrared and radio counterparts as described in Sects. 2.1 and 2.2, respectively. Table 1 gives a summary of the sensitivity and area covered by these archival data. We note that limiting the sample to galaxies at $z = 0.1$ to further minimise the impact of evolutionary trends within this redshift range leaves the measured IRRC properties unchanged within 1σ .

2.1 IR data

Here we describe the archival data underpinning our IR luminosity measurements via spectral energy distribution fitting (see Sect. 3.1.1).

2.1.1 WISE photometry

The WISE satellite (Wright et al. 2010) carried out all-sky observations in four bands, two of which (12 and $22\mu\text{m}$, with resolutions of 6.5 and 12.0 arcsec, respectively) lie in the 8–1000 μm window underpinning our total IR (TIR) luminosity measurement. Rapidly changing SED amplitudes due to polycyclic aromatic hydrocarbon (PAH) features around $12\mu\text{m}$ make modelling sources with photometric redshifts

³ For conversion factors between our assumption and other widely-used IMF models see e.g. Madau & Dickinson (2014)

⁴ We restrict ourselves to high-quality photometric redshifts by imposing the criteria `photoErrorClass = 1`, `nnCount > 95`, and $0 < z_{\text{Err}} < 0.03$. The SDSS spectroscopic and photometric catalogues were combined using the `fluxID`, which – for each spectroscopic source – identifies the corresponding SDSS photometric object (`objID`) that contributes most to the spectrum.

Table 1. Sensitivities and sky coverage of surveys used to construct our catalogue. The 5σ depths quoted for the PACS and SPIRE Point Source Catalogs (PPSC and SPSC, respectively) represent their median tabulated 5σ flux uncertainties. PPSC $100\mu\text{m}$ depth, marked by (*) in the table, is most likely underestimated, due to relatively low number of detections permitting a statistically less robust noise simulation. We increased the nominal 1 mJy FIRST detection limit by 22% to reflect the scaling we applied to FIRST fluxes in our sample to compensate missing large scale flux in our low- z sample (see Sect. 2.2). All other values are taken from the corresponding data release papers cited in the text.

	5σ [mJy]	area [deg ²]
FIRST	1.22	10,575
NVSS	2.5	37,216
IRAS $60\mu\text{m}$	200	full sky
IRAS $100\mu\text{m}$	1000	full sky
H-ATLAS DR1 $100\mu\text{m}$	220	161
H-ATLAS DR1 $160\mu\text{m}$	245	161
H-ATLAS DR1 $250\mu\text{m}$	37	161
H-ATLAS DR1 $350\mu\text{m}$	47	161
H-ATLAS DR1 $500\mu\text{m}$	51	161
PPSC $100\mu\text{m}$	107*	~3,300
PPSC $160\mu\text{m}$	236	~3,300
SPSC $250\mu\text{m}$	73	~3,700
SPSC $350\mu\text{m}$	73	~3,700
SPSC $500\mu\text{m}$	78	~3,700
WISE $22\mu\text{m}$	4.35	full sky

difficult. To avoid these issues, and to also further minimize the impact of any residual contamination from mid-IR (MIR) torus emission from AGN hosts which were not picked up by our AGN removal criteria in Sect. 3.2, we only use $22\mu\text{m}$ flux densities in the following.

Exploiting the high-resolution SDSS optical data, Lang et al. (2016) performed flux extraction with a forced photometry approach on un-blurred, co-added WISE images (Lang et al. 2014) at over 400 million optical source positions, resulting in the unWISE catalogue. The unWISE data are hence naturally linked to our SDSS parent sample. To enter our sample, we require each unWISE $22\mu\text{m}$ detection to have a signal-to-noise ratio (SNR) of at least 5, despite the availability of lower significance measurements due the forced photometry technique used for unWISE.

2.1.2 IRAS photometry

The Infrared Astronomical Satellite (IRAS; Neugebauer et al. 1984) mission covered the full sky at wavelengths 12, 25, 60 and $100\mu\text{m}$ with an angular resolution varying between ~ 0.5 arcmin at $12\mu\text{m}$ and ~ 2 arcmin at $100\mu\text{m}$. We drew 60 and $100\mu\text{m}$ fluxes from the Revised IRAS Faint Source Redshift Catalog (RIFSCz) of Wang et al. (2014), which contains galaxies selected at $60\mu\text{m}$ with $\text{SNR} > 5$ while covering 60% of the sky. We discarded IRAS 12 and $25\mu\text{m}$ fluxes tabulated in the RIFSCz, respectively, due to the difficulty of fitting the PAH features of the SED for sources with photometric redshifts and the availability of better quality WISE photometry at $22\mu\text{m}$.

Wang et al. (2014) performed a likelihood ratio matching technique to combine the Faint Source Catalog with the deep WISE $3.4\mu\text{m}$ data. They then cross-matched these sources with SDSS DR10 using the WISE positions and a 3 arcsec search radius. After reconciliation of SDSS DR10 and DR12 galaxy positions, we find that 17,829 (7,261) of the $60\mu\text{m}$ ($100\mu\text{m}$) RIFSCz sources are associated with an entry in our low- z SDSS DR12 parent catalogue.

2.1.3 Herschel photometry

During its nearly four years of operation, Herschel produced thousands of maps of varying depth with two cameras: the Photoconductor Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) and the Spectral and Photometric Imaging REceiver (SPIRE; Griffin et al. 2010), operating at $60\text{--}210\mu\text{m}$ and $200\text{--}670\mu\text{m}$, respectively, with angular resolutions of 5.6–11.3 arcsec and 17–35 arcsec. This resulted in a large number of data products optimized by numerous science collaborations for different purposes. Here we make use of two data bases that provide Herschel galaxy photometry, namely (i) the Herschel-ATLAS survey (H-ATLAS; Eales et al. 2010; Valiante et al. 2016), as well as (ii) the PACS Point Source Catalog (PPSC⁵; Marton et al. 2017) and SPIRE Point Source Catalog (SPSC⁶) for which all archival data, including calibration scans, were mined in a systematic and homogeneous way.

H-ATLAS Data Release 1 (Valiante et al. 2016) covers the three equatorial fields surveyed by the GAMA (Galaxy and Mass Assembly; Driver et al. 2011) spectroscopic survey. It consists of 120,230 sources detected at $250\mu\text{m}$. It contains PACS 100 and $160\mu\text{m}$ detections at > 3 SNR, and 250, 350, and $500\mu\text{m}$ photometry at > 4 SNR considering both instrumental and confusion noise. Extended sources were identified and their fluxes extracted using appropriately sized apertures (Rigby et al. 2011). The DR1 catalogue also provides an optical identification from the SDSS DR7/8 catalogue using the likelihood ratio technique (Bourne et al. 2016). We cross-correlated the positions of H-ATLAS optical counterparts classified as secure by Bourne et al. (2016) with our SDSS DR12 parent catalogue using a search radius of 1 arcsec. This resulted in 8,752 matches (a small fraction of the H-ATLAS DR1 sources due to our redshift cut at $z = 0.2$). We note that for 3 sources we found negative fluxes at $500\mu\text{m}$ in the H-ATLAS catalogue. These measurements were removed from any subsequent analysis.

In order to fully exploit the available Herschel coverage in the SDSS footprint, we also adopt PPSC and SPSC fluxes of > 3 SNR, where H-ATLAS photometry is not available. This enables us to increase the subset of galaxies with Herschel photometry in our combined sample (see Sect. 2.3) by a factor of ~ 8 . From the PPSC and SPSC we have removed flagged (edge-flag, blend-flag, warm altitude or solar system map flag for the PPSC and an additional large galaxy flag from Jarrett et al. 2003 for the SPSC) sources, in the process retaining 64% and 69% of all 100 and $160\mu\text{m}$ catalog entries, respectively, and 97% of the 250 and $350\mu\text{m}$ sources, as well

⁵ <https://doi.org/10.5270/esa-rw7rbo7>

⁶ European Space Agency, 2017, Herschel SPIRE Point Source Catalogue, Version 1.0. <https://doi.org/10.5270/esa-6gfkpzh>

as 96% of all 500 μm sources. To assign optical counterparts to these remaining PPSC and SPSC entries, we adopted a band-dependent matching radius. The details of this cross-correlation procedure and estimates of spurious match fractions are provided in Appendix A. We find that 5,878 100 μm PPSC sources have an optical counterpart in our parent catalog; at 160 μm this is the case for 10,149 sources. For the SPSC, we were able to assign 43,665 sources to an optical counterpart at 250 μm , 15,614 sources at 350 μm , and 2,806 sources at 500 μm .

Relying on point source measurements extracted with point spread functions ranging from 7 to 35 arcsec in angular size carries the risk of underestimating fluxes for our low-redshift, $z < 0.2$, galaxies. We were able to assess whether our PPSC and SPSC fluxes are subject to any systematic bias by comparing them to H-ATLAS photometry for objects where both types of measurements are available. This comparison reveals an average deficit of ~ 19 and 10% for PPSC fluxes relative to H-ATLAS measurements at 100 μm and 160 μm , while SPIRE fluxes are consistent within 5% in all bands, suggesting that resolution-related effects only noticeably bias our PPSC photometry. To mitigate these systematics we applied statistical corrections to our PPSC data and Appendix B details how we derived the appropriate scaling factors.

In a further test of the overall consistency of our Herschel photometry we also investigated the agreement of flux errors between H-ATLAS and the point source catalogues. For the PPSC flux errors, we adopt either the local RMS or the so-called structure noise⁷, whichever is larger. Both approaches account for the instrumental as well as the confusion noise. We note that in the case of 100 μm data, uncertainties are likely underestimated due to fewer detections available for producing noise simulation maps (78 in contrast to 326 at 160 μm), resulting in less accurate modelling of the noise in this band, especially for faint sources (based on priv. comm with Gábor Marton). To consider the uncertainties in the consistency between resolved H-ATLAS and point source PPSC/SPSC photometry, we increased the flux density errors in each band for PPSC/SPSC data as described in Appendix B. However, we emphasize that this was done only to inform our spectral energy distribution modelling (Sect 3.1.1). For source selection purposes we worked with the errors as tabulated in the catalogues.

2.2 Radio data

Our radio fluxes were drawn from two wide-area 1.4 GHz VLA surveys, FIRST and NVSS. FIRST focused on the SDSS footprint as established in Helfand et al. (2015), with higher resolution (5 vs. 45 arcsec) and sensitivity (1 vs. 2.5 mJy) than NVSS, while NVSS surveyed the entire Northern sky (Condon et al. 1998). These two catalogues were combined, via positional cross-match, and form the basis of the Unified Radio Catalog (Kimball & Ivezić 2008, 2014). An entry in the Unified Radio Catalog is either a FIRST object

with an NVSS match, an NVSS object with a FIRST match, or an unmatched object from either survey. The three closest matches within 30 arcsec to a FIRST or NVSS source were also recorded. As a result, several sources appear more than once in the final database. To facilitate flexible and easy sample selection, Kimball & Ivezić (2008) defined several flags indicating whether FIRST or NVSS was the primary matching source, and the number of objects from the other catalogue within 5, 10, 30 and 120 arcsec. We made use of these flags in the process of creating our catalogue, as explained in the following.

When both FIRST and NVSS data are available, FIRST does not always clearly provide the best total intrinsic radio flux density of a galaxy, despite its superior resolution and depth. Due to uv coverage from short baselines being absent in the B-array FIRST data, its sensitivity to extended emission is limited. Larger scale flux components thus were potentially captured only by NVSS. As a result, Helfand et al. (2015) report a 1–5% statistical flux deficit in the total FIRST sample compared to NVSS. The difference is expected to be larger on average for low redshift galaxies due to their larger angular sizes. On the other hand, multiple galaxies could be blended in the 45 arcsecond NVSS imaging, leading to positive flux biasing. To mitigate the effects of these on our analysis, we selected sources in the Unified Radio Catalog that are:

- (i) NVSS sources with no FIRST counterparts within the 30 arcsec of the NVSS position. Expressed with the flags of the catalogue this selection is $(\text{matchflag_nvss} = -1) \ \& \ (\text{matchflag_first} = 0) \ \& \ (\text{matchtot_30} = 0)$.
- (ii) NVSS sources with only one FIRST detection in their beam. Selection flags for these sources were $(\text{matchflag_nvss} = -1) \ \& \ (\text{matchflag_first} = 1) \ \& \ (\text{matchtot_30} = 1)$.
- (iii) FIRST sources with no NVSS counterparts, i.e. $(\text{matchflag_first} = -1) \ \& \ (\text{matchflag_nvss} = 0) \ \& \ (\text{matchtot_30} = 0)$.

This subset of the Unified Radio Catalog was then spatially matched to our low- z SDSS parent sample. For sources in (i) and (ii) we used NVSS positions and a matching radius of 30 arcsec (equal to the value adopted for the NVSS matching in Kimball & Ivezić 2008, 2014), and for sample (iii) FIRST positions with 2.5 arcsec matching radius (half of the 5 arcsec FIRST beam size). Comparing FIRST and NVSS fluxes in Fig. 3 we find a systematic NVSS/FIRST flux ratio of ~ 1.22 . This is largely independent of 1.4 GHz angular size, flux and galaxy type, i.e. SFG or AGN. However, the ratio is closer to 1 (~ 1.09) for galaxies selected via (ii) if we do not restrict the selection to low- z sources, supporting the idea that the main reason for the flux offset is missing extended flux in FIRST measurements of nearby sources (see e.g. Helfand et al. 2015). An alternative explanation for the offset is the possibility that NVSS fluxes are biased high due to blending. We consider this a less likely scenario, because we removed all NVSS sources with more than one FIRST counterpart in order to minimize the effect of blending. Thus, we conclude that for our purposes in this particular sample NVSS provides a more robust measure of the 1.4 GHz flux. As a result we adopted NVSS measurements for (i) and (ii) and applied a statistical correction of 1.22 to all FIRST detections in (iii), in order to make these fluxes consistent with (i) and (ii). This corresponds to a 0.09

⁷ The structure noise produces statistical estimates on the error of the photometry by measuring the flux of artificial sources injected into the various Herschel fields (for details see the HPPSC Explanatory Supplement and references therein).

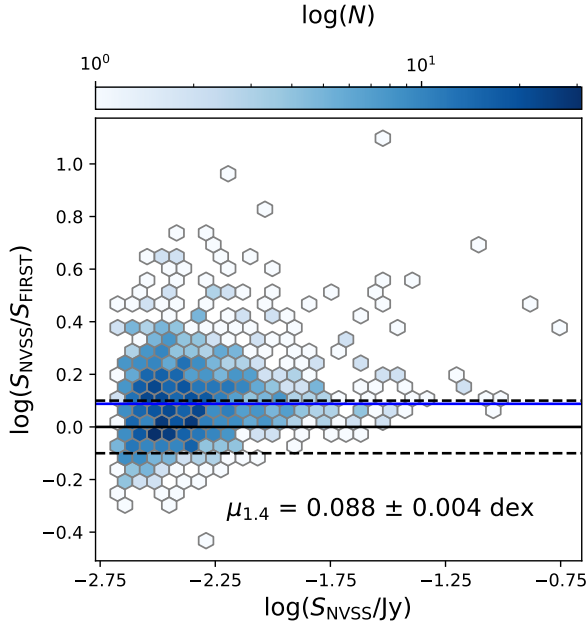


Figure 3. Logarithmic ratio of FIRST and NVSS flux densities as a function of the latter using sources in (ii) from Sect 2.2. Horizontal black line is drawn at 0, i.e. a flux density ratio of 1, while the dashed lines represent the ± 0.1 dex offsets. The median ratio, $\mu_{1.4}$, shown as a horizontal blue line, was used to scale the FIRST flux densities to statistically match the NVSS data.

dex upward correction in logarithmic 1.4 GHz flux space (i.e. a 0.09 dex downward shift of the IRRC parameter, q) for these objects. Table 2 gives a summary of source counts in (i), (ii) and (iii) in our main sample and its subsamples (defined below).

We note that the final FIRST catalog release (Helfand et al. 2015) was published after the assembly of the Kimball & Ivezić (2014) Unified Radio Catalog. The Helfand et al. (2015) catalog contains significantly more robust sidelobe probability estimates, updated flux measurements compared to those in (Kimball & Ivezić 2014) and excludes data from unreliable FIRST pointings. We incorporated these improvements by matching the final FIRST catalogue to the FIRST sources of the Unified Radio Catalog via a simple positional cross-match using a matching radius⁸ of 1 arcsec. We also removed FIRST sources with a sidelobe probability greater than 10% to mitigate contamination by spurious detections prior to the selection steps described above.

2.3 Combined IR and radio sample

To summarize, we have collected IR and radio flux densities at 8 different wavelengths using 6 archival databases with varying depths and survey areas, as presented in Table 1. A

⁸ The description of the latest FIRST catalogue – available at <http://sundog.stsci.edu/first/catalogs/readme.html> – suggests a better than 1 arcsec positional accuracy at the detection limit of the survey, and 0.5 arcsec for $\sim 10\sigma$ detections. Accordingly, after the cross-match we find a mean separation of 0.01 arcsec with a scatter of 0.06 arcsec.

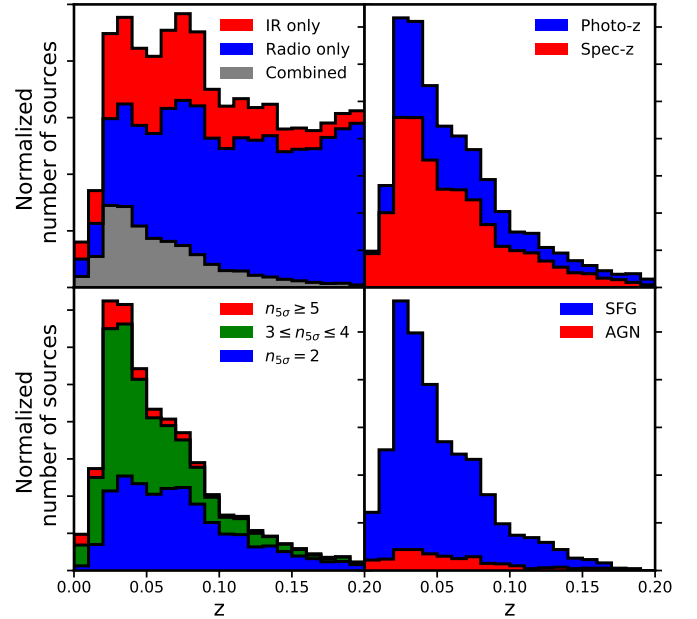


Figure 4. Stacked histograms showing the key sample properties. **(top left)** – redshift distributions of the only radio- and IR-detected samples, and the combined sample, i.e. sources detected both at IR and radio wavelengths; **(top right)** – sources in the combined sample with photometric and spectroscopic redshifts; **(bottom left)** – sources in the combined sample with fluxes detected at the $\geq 5\sigma$ level in 2, 3 or 4, and ≥ 5 IR bands; **(bottom right)** – galaxies classified as AGN (both optical and WISE selected) or star-forming galaxies in the combined sample.

source is required to have a $\text{SNR} > 5$ detection in the unWISE catalogue at $22\mu\text{m}$ and at least one other $\text{SNR} > 5$ measurement in any of the longer wavelength data in Sect 2.1 to enter our IR-selected sample. Meanwhile, to be considered as a radio-detected object, each source needed an at least 5 SNR 1.4 GHz flux density measurement in the (i), (ii) or (iii) subsamples of the Unified Radio Catalog, as described in Sect 2.2. Our compilation thus lead to samples of 25,782 IR- and 51,774 radio-detected galaxies at $z < 0.2$, respectively. Merging these IR- and radio-detected catalogues resulted in a joint catalogue of 67,908 objects. In this joint sample, 9,645 sources are members of both the radio- and IR-detected catalogues and will be referred to as combined sample henceforth, while 16,134 have only IR and 42,126 only radio data. In Fig. 4 we show the redshift distributions of IR- and radio-detected sources and the combined sample; sources with spectroscopic and photometric redshifts in the combined sample; IR-detected sources with different number of IR photometric bands available; and SFGs and AGN. For the latter, Sect. 3.2 details our classification approach.

Table 2 contains the number of detections in the combined sample, as well as other samples introduced later in Sect. 3, per photometric band. Fig. 5. shows the IR flux densities in each band against the 1.4 GHz radio flux densities in our combined sample. The IR-radio correlation is apparent already before the conversion to luminosity and removal of contaminating AGN sources. Due to the mismatch between sensitivities of the IR and radio, a selection bias affects the total and the monochromatic IRRC statistics. To mitigate

Table 2. Number of galaxies detected in the different IR bands in the combined and the depth-matched samples, as well as for the SFG and AGN subsamples drawn from the latter. Bracketed percentages are the fraction of source detected at $>5\sigma$ in any given band and sample (e.g. of the 972 galaxies in the combined sample with a measured Herschel/SPIRE 500 μm flux, 90% – i.e. 875 objects – have a $\text{SNR} > 5\sigma$ detection). The final three rows list the number of sources with only NVSS fluxes, with both NVSS and FIRST fluxes, and with only FIRST detections, respectively.

	Combined	Depth-matched	Depth-matched SFG	Depth-matched AGN
unWISE 22 μm	9,645 (100%)	6,601 (100%)	2,371 (100 %)	248 (100 %)
IRAS 60 μm	8,720 (90%)	6,217 (94%)	2,258 (95%)	229 (92 %)
IRAS 100 μm	4,407 (74%)	3,519 (77%)	1,338 (77%)	122 (73 %)
Herschel 100 μm	442 (92 %)	299 (95%)	111 (97%)	9 (81 %)
Herschel 160 μm	415 (91%)	263 (95%)	98 (96%)	11 (100 %)
Herschel 250 μm	1,710 (99%)	1,067 (99%)	347 (99%)	38 (100 %)
Herschel 350 μm	1,683 (98%)	1,058 (99%)	354 (99%)	37 (97 %)
Herschel 500 μm	914 (89%)	697 (93%)	262 (95%)	24 (89 %)
NVSS only	3,117	1,280	597	36
NVSS+FIRST	4,442	3,891	1,289	166
FIRST only	2,089	1,430	485	46

this, we derived flux density cuts that already appear in Fig. 5. However, before we detail how these were calculated in Sect. 3.3, we first describe our luminosity estimation approach in Sect. 3.1.

3 METHODS

In this section we describe our approach to calculating IR and radio luminosities for the 9,645 objects in the combined sample (Sect 2.3), and then classifying these as AGN or SFG galaxies. We define a depth-matched subsample of the combined sample in Sect. 3.3 in an effort to mitigate the effect of the sensitivity mismatch between our radio and IR catalogues (Table 1) on the IRRC properties studied in Sect. 4.

3.1 Infrared and radio luminosity derivations

3.1.1 IR luminosities from SED fitting

In order to estimate the total (8 – 1000 μm) infrared luminosity (L_{TIR})⁹ of our galaxies, we fitted their IR flux densities with the SED templates of Dale & Helou (2002) (see also Dale et al. 2001)¹⁰. This SED library contains IR spectra of different shapes, sorted by their radiation field hardness parameter. We assigned total infrared luminosity values to each template following the relation between S_{60}/S_{100} and L_{TIR} from Marcellac et al. (2006). We then sorted the library SED templates according to their IR luminosity, normalised them, and carried out a cubic spline interpolation between the spectra. This allows us to draw not only one of

the 119 pre-defined spectra from Dale & Helou (2002), but transitional shapes between them from the interpolation via a continuous shape parameter, γ . To relax the assumption made when we ordered the SED library, i.e. that SED shape is intrinsically tied to the IR luminosity of the sources, we fitted γ and L_{TIR} independently of each other by maximizing the logarithmic likelihood function

$$\mathcal{L}(L_{\text{TIR}}, \gamma) = -0.5 \sum_{i=1}^N \left(\frac{\log(S_i) - \log(f(\lambda_i, L_{\text{TIR}}, \gamma))}{\sigma_i} \right)^2 + \mathcal{P}, \quad (1)$$

where S_i and σ_i are the observed flux density and its logarithmic uncertainty¹¹ in photometric band i . $f(\lambda_i, L_{\text{TIR}}, \gamma)$ is the flux density in the same band predicted by the best-fit model with parameters L_{TIR} and γ considering the bandpass shape of band i , and

$$\mathcal{P} = \begin{cases} 0, & \text{if } 10^5 \leq L_{\text{TIR}} \leq 10^{14} \text{ and } 0 \leq \gamma \leq 119 \\ -\infty, & \text{otherwise} \end{cases} \quad (2)$$

is the so-called prior function. \mathcal{P} ensures our optimization process probes a physically meaningful parameter space and γ is interpolated between the 119 SED templates.

To find the best-fit SED model with realistic error estimates on the free parameters, L_{TIR} and γ , we used the affine invariant Markov chain Monte Carlo (MCMC) sampler EMCEE (Foreman-Mackey et al. 2013), a free, open-source code implemented in PYTHON. We initiated 100 walkers with L_{TIR} values randomly generated according to a uniform distribution between the limits of Eq 2, the prior function. In combination with our choice of using $\log(S_i)$ over S_i in the likelihood function (eq. 1), this setup resulted in the fastest and most robust convergence of the likelihood sampling process. Specifically, we found that the burn-in period of a typical MCMC chain is ~ 200 steps. Thus, to achieve sufficient sampling we ran our walkers for 2,200 steps and produced

⁹ In recent years it has become common practice to denote luminosity in the 8 – 1000 μm range as L_{IR} , however, to clearly distinguish between the total infrared and far-infrared radio correlations, we chose this notation throughout the paper.

¹⁰ We note that, on average, L_{TIR} values obtained from SED fits using the Chary & Elbaz (2001) template library are consistent with the ones produced by the Dale & Helou (2002) library within 0.01 dex. We decided to use the latter, because it yielded overall better quality fits and smoother L_{TIR} posterior distributions.

¹¹ We approximated logarithmic flux density errors as $\sigma_i = 0.434 \cdot \Delta S_i / S_i$, where ΔS_i is the uncertainty of the flux density in photometric band i .

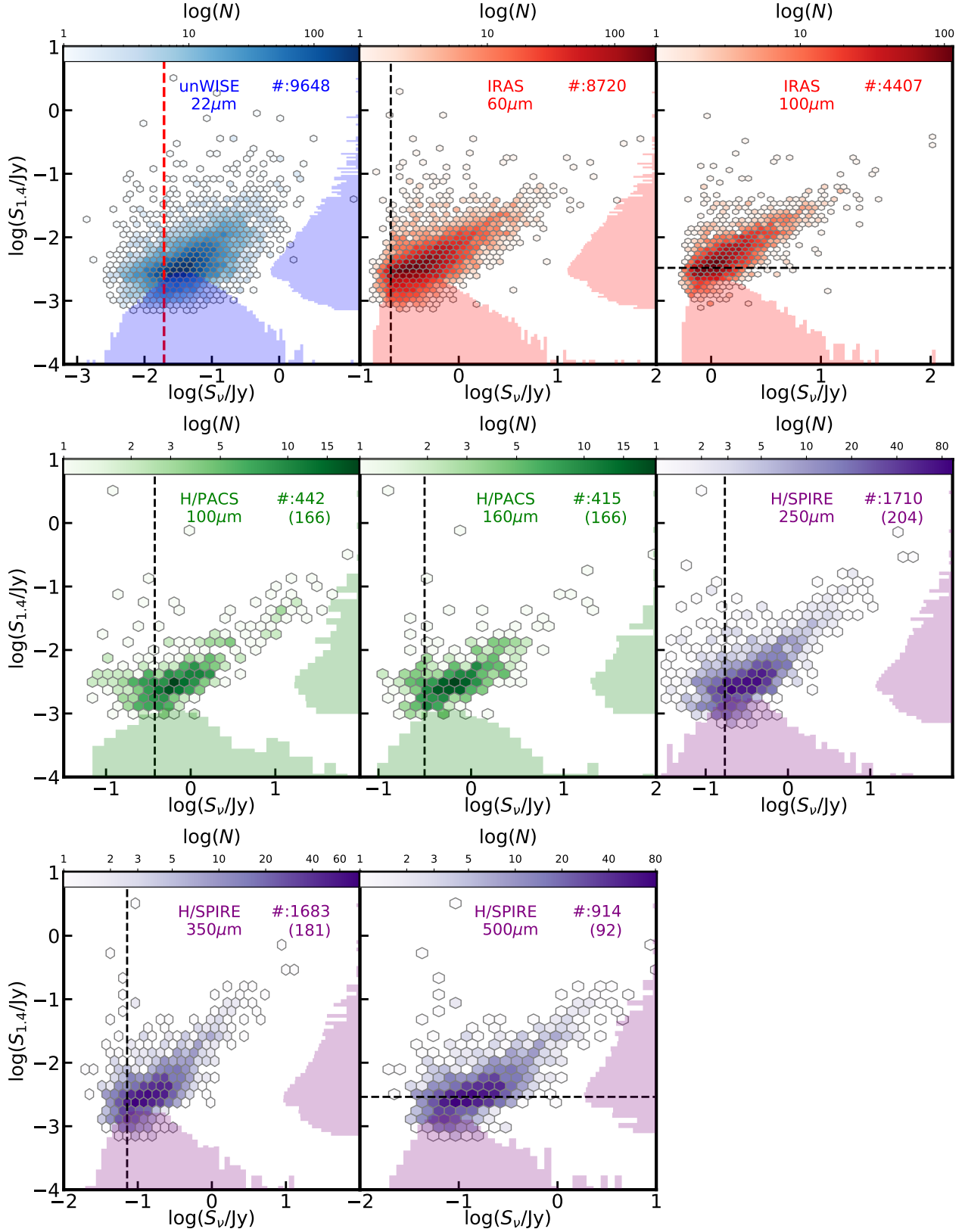


Figure 5. Radio versus infrared fluxes from unWISE (22 μm), IRAS (60 and 100 μm) and Herschel (100, 160, 250, 350 and 500 μm) in our joint sample (i.e. before flux cuts and removing AGN and ambiguous galaxies). IR and radio flux density distributions are shown on the sides. The dashed lines in each panel illustrate the flux density cuts we applied in order to select our radio-IR depth-matched samples for measuring monochromatic IRRC properties, as described in Sect. 3.3. Table 4 contains the corresponding flux density values. We note that only the ~ 19.5 mJy flux selection at 22 μm was used for the depth-matched sample for investigating the total IRRC. In the upper right corner of each panel we display the number of sources with both radio and IR-detections at a given wavelength. For Herschel bands, values in brackets denote the size of the subset of galaxies with fluxes from the H-ATLAS catalog.

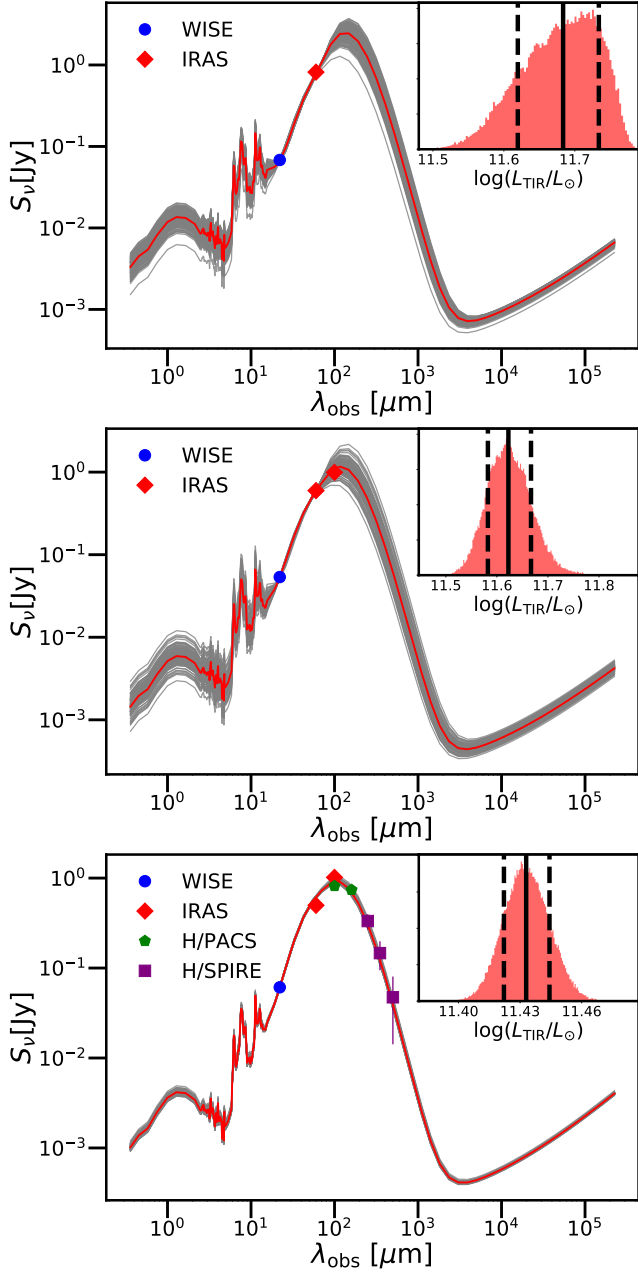


Figure 6. Examples of fitted spectral energy distributions with different photometric coverage (i.e. 2, 3 and 6 IR detections of at least 5 SNR from top to bottom), using the templates of Dale & Helou (2002). Red lines are the best-fit curves, grey ones are 100 randomly selected models from converged MCMC chains, representing the $\sim 1\sigma$ confidence intervals of each fit. The insets show the L_{TIR} posterior distributions, with the medians (16th/84th percentiles) highlighted by solid (dashed) black lines.

marginalised distributions and statistics after removing the first 200 samples from each, while they were converging on the best-fit parameters. A median acceptance rate¹² of 0.4

¹² Acceptance rate or fraction allows for a quick check on MCMC convergence. It is defined as the fraction of proposed steps that are accepted in the chain. An acceptance rate of ~ 0 indicates that

with a 0.1 standard deviation indicates that the majority of our fits indeed sufficiently converged.

Finally, we assessed the quality of each SED model in order to identify and remove poor SED fits. Even though χ^2_{red} is a widely used metric to judge the goodness of a model fit, it has some potential pitfalls as noted by, e.g., Andrae et al. (2010). Chiefly, the number of degrees of freedom for non-linear models (such as our SED templates containing black-body radiation curves of different temperatures), in general, can be anywhere between 0 and $N-1$, where N is the number of data points, and may even change during the fit, rendering the use of a single χ^2_{red} value cut to separate poor and robust models inadequate across all our sources which have varying numbers of available photometric bands. To circumvent this issue, we examined the normalised logarithmic residuals, R_i , for each source, defined as:

$$R_i = \frac{\log(S_i) - \log(f(\lambda_i, L_{\text{TIR}}, \gamma))}{\sigma_i}. \quad (3)$$

Given the data, the true model should produce normalised residuals that follow a standard normal distribution. There are a wide variety of tests to assess whether a set of data is likely to be drawn from such a distribution. However, in order to be sufficiently robust, these require samples larger than the maximum number of 8 measurements an individual galaxy in our analysis can have. Therefore, as a simpler approach, we computed the mean of the normalized residuals, μ_0 , for our sources with N photometric datapoints as

$$\mu_0 = \frac{1}{N} \sum_{i=0}^N R_i. \quad (4)$$

Since μ_0 can be interpreted as the mean offset between the model and the data in signal-to-noise space, we flagged models as of poor quality, if μ_0 deviated from 0 by more than 1, i.e. on average our model was not consistent with the data within 1σ . We note that inspecting band-by-band normalised residual distributions, we found that on average all of them are consistent with 0 within 1σ and show no correlation with wavelength, suggesting that there are no statistically significant systematic errors with the flux measurements and that our model library covers the observed IR colour space.

With this method we identified 1,989 sources with fitted models inconsistent with the data, $\sim 21\%$ of our overall combined sample. Among star forming galaxies (see Sect. 3.2 for details), 386 (14%) proved to have poor fits. These were excluded in the subsequent analysis.

Fig. 6 shows typical SEDs with best-fit model and additional randomly drawn models from the MCMC chains representing the 1σ confidence interval of our fit alongside

almost all proposed steps are rejected, the chain essentially being stuck and generating very few independent values, such that it does not properly sample the posterior distribution. Vice versa, an acceptance rate of ~ 1 means that nearly all new steps are accepted, effectively resulting in a random walk, which also does not probe the sought probability density distribution. Depending on the number of free parameters, values between ~ 0.2 and 0.5 are often considered a sign of a well-sampled posterior distribution.

the marginalised L_{TIR} distribution from the posterior sampling, which was used to derive L_{TIR} uncertainties. Typical L_{TIR} errors in our depth-matched catalogue are 0.12, 0.05 and 0.02 dex for sources with 2, 3–4, and >4 available photometric bands, respectively.

The fitted SEDs were used to derive an empirical K-correction at various wavelengths by taking the ratio of the observed and rest-frame fluxes for each source (see Sect. 4.1 for more details). These corrections were applied to the closest adjacent flux measurement when computing the monochromatic IR luminosities presented in Fig. 8. Finally, for a more direct comparison with e.g. Yun et al. (2001), we also calculated the far-IR luminosity, L_{FIR} , for each galaxy by integrating our best-fit SED models between 42.5 – 122.5 μm .

3.1.2 1.4 GHz radio luminosity estimates

Radio flux densities were converted into 1.4 GHz rest-frame radio continuum luminosities, $L_{1.4}$, using

$$\left(\frac{L_{1.4}}{\text{WHz}^{-1}}\right) = C_1 \frac{4\pi}{(1+z)^{(1+\alpha)}} \left(\frac{D_L}{\text{Mpc}}\right)^2 \left(\frac{S_{1.4}}{\text{mJy}}\right), \quad (5)$$

where $C_1 = 9.52 \times 10^{15}$ is the conversion factor from $\text{Mpc}^2 \text{mJy}$ to WHz^{-1} , α is the radio spectral index¹³, z is redshift, D_L is the luminosity distance and $S_{1.4}$ is the measured 1.4 GHz flux density. We used the typical $\alpha = -0.7$ assumption (e.g. Kimball & Ivezić 2008).

3.2 Identification and removal of AGN host galaxies

It is generally assumed that the IRRC emerges from the correlation of IR and radio flux densities with star formation activity. Since our primary aim is to study this relation, we selected sources in our sample identified as predominantly star-forming. On the other hand, since the IRRC can also be used to detect excess radio emission presumably linked to AGN activity, we also investigated AGN in our sample.

3.2.1 Classification based on optical emission lines

For galaxies with SDSS DR 8 emission line measurements from the value-added MPA/JHU group¹⁴, we classify galaxies as star-forming, AGN or composite following the method presented in Kewley et al. (2006) (and first introduced by Baldwin et al. 1981). The method makes use of the emission line ratios $[\text{NII}]\lambda 6584/\text{H}\alpha$, $[\text{SII}]\lambda 6717, 6731/\text{H}\alpha$, $[\text{OI}]\lambda 6300/\text{H}\alpha$, and $[\text{OIII}]\lambda 5007/\text{H}\beta$ which are sensitive to the metallicity and ionization properties of the gas. First, we use all the diagnostic line ratios to classify star-forming galaxies using the theoretical “maximal starburst line” derived by Kewley et al. (2001), indicating the theoretical maximum line ratios that could be produced by pure stellar photoionization models alone.

¹³ The radio spectral index is defined as $S_\nu \propto \nu^\alpha$, where S_ν is the flux density at frequency ν .

¹⁴ http://www.sdss3.org/dr10/spectro/galaxy_mpa_jhu.php

We then identify galaxies which are classified as star-forming according to their $[\text{SII}]/\text{H}\alpha$ and $[\text{OI}]/\text{H}\alpha$ ratios, but fall in the “composite” region in the $[\text{NII}]/\text{H}\alpha$ vs. $[\text{OIII}]/\text{H}\beta$ diagnostic according to the empirical Kauffmann et al. (2003) boundary. These are galaxies with a composite spectrum containing a mix of HII region emission and a harder ionizing source, and are thus not included in our SFG sample.

We classify as AGN the sources that lie above the “maximal starburst line” in the $[\text{SII}]/\text{H}\alpha$ and $[\text{OI}]/\text{H}\alpha$ line ratio diagnostics and above the Kauffmann et al. (2003) line in the $[\text{NII}]/\text{H}\alpha$ diagnostic diagram.

Sources we label as unclassified lack sufficiently high significance line measurements (we require $\text{SNR} > 3$ in $\text{H}\alpha$ following Leslie et al. (2016)) or observed spectra altogether. Even though it is possible to utilize e.g. a colour-colour selection of SFGs and AGN to increase our sample size, the unclassified population has both radio and IR luminosity distributions coincident with the classified ones, and thus by relying solely on emission lines we retain the ability to probe the entire luminosity regime available in our combined sample.

3.2.2 Identifying AGN using mid-IR colours

Assef et al. (2018) selected AGN candidates using the WISE 3.4 and 4.6 μm bands. The selection criteria were calibrated and assessed based on UV- to near-IR spectral energy distribution analysis of AGN in the NOAO Deep Wide-Field Survey Boötes field where deep WISE data are available. We use their criteria that select AGN candidates with a 90% confidence. Galaxies meeting the following criteria are flagged as MIR-selected AGN, where W1 and W2 are the WISE 3.4 and 4.6 μm Vega magnitudes, SNR is SNR, and cflags = 0¹⁵ for W1 and W2:

$$\begin{aligned} W1 &> 8, \\ W2 &> 7, \\ \text{SNR}_{W2} &> 5, \\ W1 - W2 &> \begin{cases} a \exp[b(W2 - c^2)] & W2 > c \\ a & W2 < c, \end{cases} \end{aligned}$$

where $(a, b, c) = (0.650, 0.154, 13.86)$. This selection identifies 4,470 MIR-AGN candidates in the joint catalogue (5%) and these AGN candidates tend to be more luminous than the optically selected AGN candidates. Only 828 (186) objects in the IR-detected (combined IR and radio detected) catalogue satisfy both the MIR and optical AGN selection criteria, indicating the importance of a multiwavelength approach for selecting all types of AGN.

As described in Sect. 3.1.1, the templates used for IR SED fitting assume that the IR emission arises purely from star formation. If a source has a non-negligible AGN-related MIR component, which in most cases enhances only the MIR flux but not the FIR ones, our fitted L_{TIR} values are very

¹⁵ As described in the Explanatory Supplement to WISE (<https://wise2.ipac.caltech.edu/docs/release/allsky/expsup>), cflag stands for contamination and confusion flag. It indicates whether a source may be affected by a nearby imaging artifact. A value of 0 indicates it is not.

Table 3. Number of SFG, AGN and composite sources classified by optical emission lines (Sect. 3.2.1) in the combined sample and its subset of depth-matched galaxies. Bracketed numbers correspond to sources identified as MIR AGN in each sample (Sect. 3.2.2). These were excluded from our analysis due to their most likely inaccurate L_{TIR} estimates. Unclassified galaxies lack sufficiently high quality spectra that permits classification.

	Combined	Depth-matched
SFG	2,495 (54)	2,093 (46)
AGN	313 (80)	239 (67)
Composite	1,417 (77)	1,107 (68)
Unclassified	3,431 (293)	1,743 (190)
Total	7,656 (504)	5,182 (371)

likely to be overestimated. We therefore excluded all 504 AGN identified at MIR wavelengths. This leaves 2,441 SFGs and 233 AGN in our sample that show no sign of AGN activity at MIR wavelengths, and have reliable L_{TIR} estimates according to our residual analysis described in Sect. 3.1.1.

3.3 Depth homogenization and depth-matched sample

Up to now our selection strategy has produced a sample of 7,656 jointly IR- and radio-detected $z < 0.2$ galaxies with robust SED fits, of which 2,441 were identified as high-confidence SFGs in the preceding section. We now consider the bias affecting IRRC statistics arising from sensitivity differences between IR and radio (see, e.g., Sargent et al. 2010a). To assess the possibility of such an effect on our study, we compared the depths of the data sets used to create our catalogue.

Fig. 7 shows the total IR and radio luminosity limits as a function of redshift estimated for the various photometric bands and surveys in our catalogue. For each IR band/survey, at a given redshift, we calculate the predicted flux density value of every SED template in the Dale & Helou (2002) library taking the transmission curves of the specific instrument into account and plot the L_{TIR} value of the SED template that reproduces the 5σ flux density limit of the respective catalogue. This approach relies on the assumption that the IR colour – L_{TIR} relation in the template library holds. Each individual survey is sensitive to galaxies above its corresponding curve in Fig. 7. For radio surveys FIRST and NVSS, we calculate their radio luminosity sensitivity curves by substituting their flux detection limits into Eq. 5 in our redshift range. These are then subsequently converted into L_{TIR} by re-scaling them with the currently widely adopted value of the local $q_{\text{TIR}} = 2.64$ of Bell (2003).

Barring the $250\mu\text{m}$ H-ATLAS coverage, which contributes only $\sim 2\%$ to the footprint of our sample, unWISE $22\mu\text{m}$ observations provide the deepest data in our catalogue. The 100, 160 and $350\mu\text{m}$ H-ATLAS Herschel data are matched quite well in sensitivity with FIRST. Lastly, NVSS, IRAS and H-ATLAS $500\mu\text{m}$ provide shallower measurements relative to FIRST. NVSS and IRAS encompass a larger area and probe similar luminosity regimes in $z < 0.2$ galaxies. However, the 2 Jy IRAS $60\mu\text{m}$ flux cut imposed by Yun et al. (2001) results in significantly shallower IR coverage than the radio data from NVSS in their sample.

The consequences of such a mismatch between IR and radio measurements are explored in Sect. 4.3.

Following Sargent et al. (2010a), to avoid selection effects biasing our IRRC measurement and ensure that both our radio and IR data have comparable depths, we applied a flux cut of 19.5 mJy to our unWISE $22\mu\text{m}$ detections. This flux limit was calculated using the ratio between unWISE and FIRST luminosity limits (Fig. 7) at $z = 0.046$, the median redshift of our combined sample. However, this simple flux cut leaves 1,612 galaxies in the catalogue that are outside of the FIRST footprint, and are only covered by the shallower NVSS. These sources would require a ~ 0.4 dex higher $22\mu\text{m}$ flux cut in order to match IR and radio sensitivity in this region. In principle, we could define a second sensitivity tier in our catalogue and include these objects, however, firstly, only $\sim 5\%$ of them have high quality spectroscopy enabling an SFG/AGN classification based on spectra, and secondly, their overall IR and radio luminosity distribution closely match the luminosity range of the rest of our sample, and thus their inclusion would not improve our ability to probe the IRRC. Therefore, we simply discard all sources that were not covered by the FIRST survey when defining our final, depth-matched sample which includes 6,611 galaxies. The effects of unmatched radio and IR luminosity limits are further investigated in Sect. 4.3. The number of sources available in each photometric band across all wavelengths in the combined and the depth-matched catalogues is shown in Table 2.

Finally, to measure the unbiased monochromatic IRRC in each of our available IR bands, we introduce band-by-band flux cuts to apply to either IR or radio fluxes, depending on which is deeper in relative terms. This was done via the same method as outlined above for the $22\mu\text{m}$ data, using the median redshift of sources jointly detected in the radio and in a given IR band. Due to the heterogeneous nature of the fields used for the PPSC and SPSC catalogues it is not straightforward to quantify their overall sensitivity. We thus limit ourselves to H-ATLAS data when constructing the monochromatic IRRCs for the different Herschel bands (i.e. we do not consider galaxies which only have PPSC or SPSC flux measurements). We report all adopted band-by-band flux cuts in Table 4, and show them in Fig. 5. We note that IRAS $100\mu\text{m}$ and Herschel $500\mu\text{m}$ data, unlike all other IR bands, are less sensitive than the radio coverage, such that we have to apply a cut to the 1.4 GHz flux distribution, rather than to the 100 or $500\mu\text{m}$ fluxes. These monochromatic depth-matched samples were only used when examining monochromatic IRRCs. For all other aspects of our analysis, when referring to the “depth-matched sample”, we mean the subset of the combined sample which was selected via a single flux cut in the selection band at $22\mu\text{m}$.

4 RESULTS

4.1 Monochromatic infrared-radio correlations

For the $>5\sigma$ detections in each IR band we calculate monochromatic luminosities, νL_ν , at the rest-frame frequency ν (with ν corresponding to the characteristic wavelength/frequency of each band). Measured flux densities were converted to rest-frame monochromatic luminosities via:

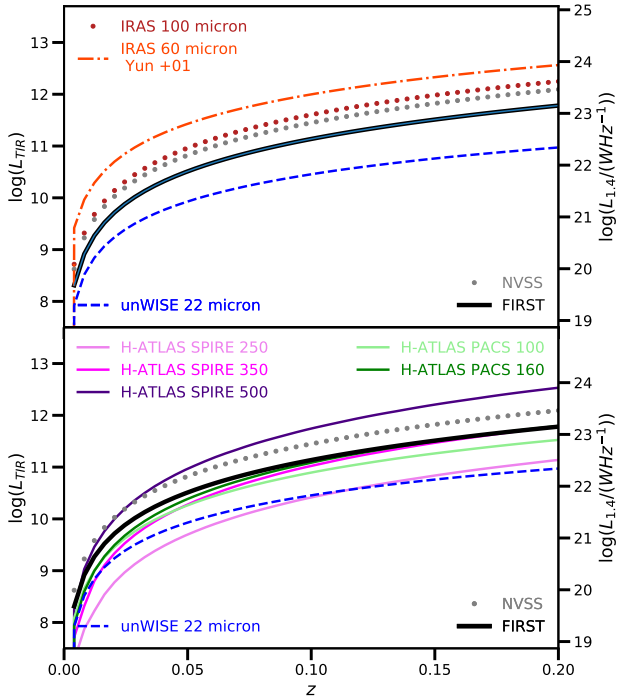


Figure 7. Redshift-dependent L_{TIR} sensitivity for the data sets used in this work. Radio luminosity limits were converted to L_{TIR} with the canonical $q_{\text{TIR}} = 2.64$ of Bell (2003). IRAS (H-ATLAS) sensitivity curves are displayed in the upper (lower) panel. For easier reference, FIRST, NVSS and WISE 22 μm limits are shown in both panels.

Table 4. Selection cuts in each IR band calculated at the median redshift, \bar{z} , of galaxies detected in the appropriate band. The WISE 22 μm flux cut was used when defining our depth-matched sample. All other cuts are instead applied exclusively when fitting monochromatic IRRCs. Both the IRAS 100 μm and the H-ATLAS 500 μm data are IR-limited, i.e. the radio data are deeper than the IR, and thus require a flux density cut at 1.4 GHz, as indicated by the third column of the table.

IR band	flux limit [mJy]	\bar{z}	flux cut applied to
WISE 22 μm	19.5	0.046	IR
IRAS 60 μm	233.6	0.044	IR
IRAS 100 μm	3.5	0.035	radio
H-ATLAS 100 μm	419.9	0.050	IR
H-ATLAS 160 μm	313.6	0.049	IR
H-ATLAS 250 μm	165.5	0.051	IR
H-ATLAS 350 μm	41.6	0.049	IR
H-ATLAS 500 μm	4.3	0.032	radio

$$\left(\frac{\nu L_\nu}{L_\odot}\right) = C_2 4\pi \left(\frac{D_L}{\text{Mpc}}\right)^2 K \left(\frac{S_{\nu,\text{obs}}}{\text{Jy}}\right) \left(\frac{\nu}{\text{Hz}}\right), \quad (6)$$

where $C_2 = 3.64 \cdot 10^{-7}$ is the conversion factor from $\text{Mpc}^2 \text{Jy Hz}$ to solar luminosity and K is the K-correction factor containing both colour correction (computed as the ratio of S_ν and $S_{\nu,\text{obs}}$, i.e. flux density at the observed wavelength) and bandpass compression terms. Uncertainties on the adopted K -correction values were calculated by re-sampling SEDs from the posterior distributions of γ . Due to the low redshift of our sample, these proved to be negligible compared to the uncertainty of the fluxes, which thus dominate the error budget of the νL_ν measurements.

As Fig. 8 shows, all observed monochromatic IR luminosities correlate with 1.4 GHz radio continuum luminosity. Each panel contains only SFG sources from samples depth-matched on a band-by-band basis (see Table 4 for the selection criteria). Arguably, it is not appropriate to treat either the radio or the IR luminosities as the independent variable. Correspondingly, Bell (2003) carried out a bisector fit to determine the slope of the relation. However, as Hogg et al. (2010) pointed out, it is preferable to adopt other approaches for linear regression. Therefore, we inferred the best-fit model

$$\log\left(\frac{L_{1.4}}{\text{WHz}^{-1}}\right) = m \cdot \log\left(\frac{\nu L_\nu}{L_\odot}\right) + b \quad (7)$$

with the bivariate correlated errors and intrinsic scatter (BCES; Akritas & Bershady 1996; Nemmen et al. 2012) method, in particular by minimizing the squared orthogonal distances to the modelled relation. We measure the dispersion as the standard deviation of the orthogonal offset distribution of the data relative to the best-fit model.

Table 5 contains the slopes (m_{all} , m_{dm} and m_{dmSFG}), intercepts (b_{all} , b_{dm} and b_{dmSFG}) and dispersions (σ_{all} , σ_{dm} and σ_{dmSFG}) of all monochromatic IRRCs for the full combined sample, the depth-matched sample (see Sect. 4.3) and its subset of depth-matched SFGs, only using galaxies with at least 5σ flux density measurements in a given IR band. Removing AGN from the samples reduces the dispersion of all correlations, in some cases by almost 50%, and in particular for Herschel data it systematically brings the IIRC slope closer to unity. This is due to the radio-loud AGNs that tend to be high radio luminosity outliers and are typically IR luminous as well, thus simultaneously steepening the IIRC slope and adding to its dispersion. We include the fits to the full combined sample in order to illustrate the effect our flux-matching approach has on the derived monochromatic IIRC parameters. The most striking example of this bias occurs when we compare the 100 μm IRRCs based on IRAS and H-ATLAS data. Initially, with no flux cut applied, their slopes are inconsistent at the $\sim 2.5\sigma$ level, but fitting their depth-matched SFG subsamples we find IIRC parameters consistent within 1σ .

Overall, as shown in Fig. 9, we find slopes (m_λ) near unity below 100 μm transitioning to slopes of ~ 1.2 in the 160 – 350 μm regime. This change of IIRC monochromatic slope values likely evidences the transition from bands probing warmer dust emission (which more closely correlates with on-going star-formation activity) to a regime sampling

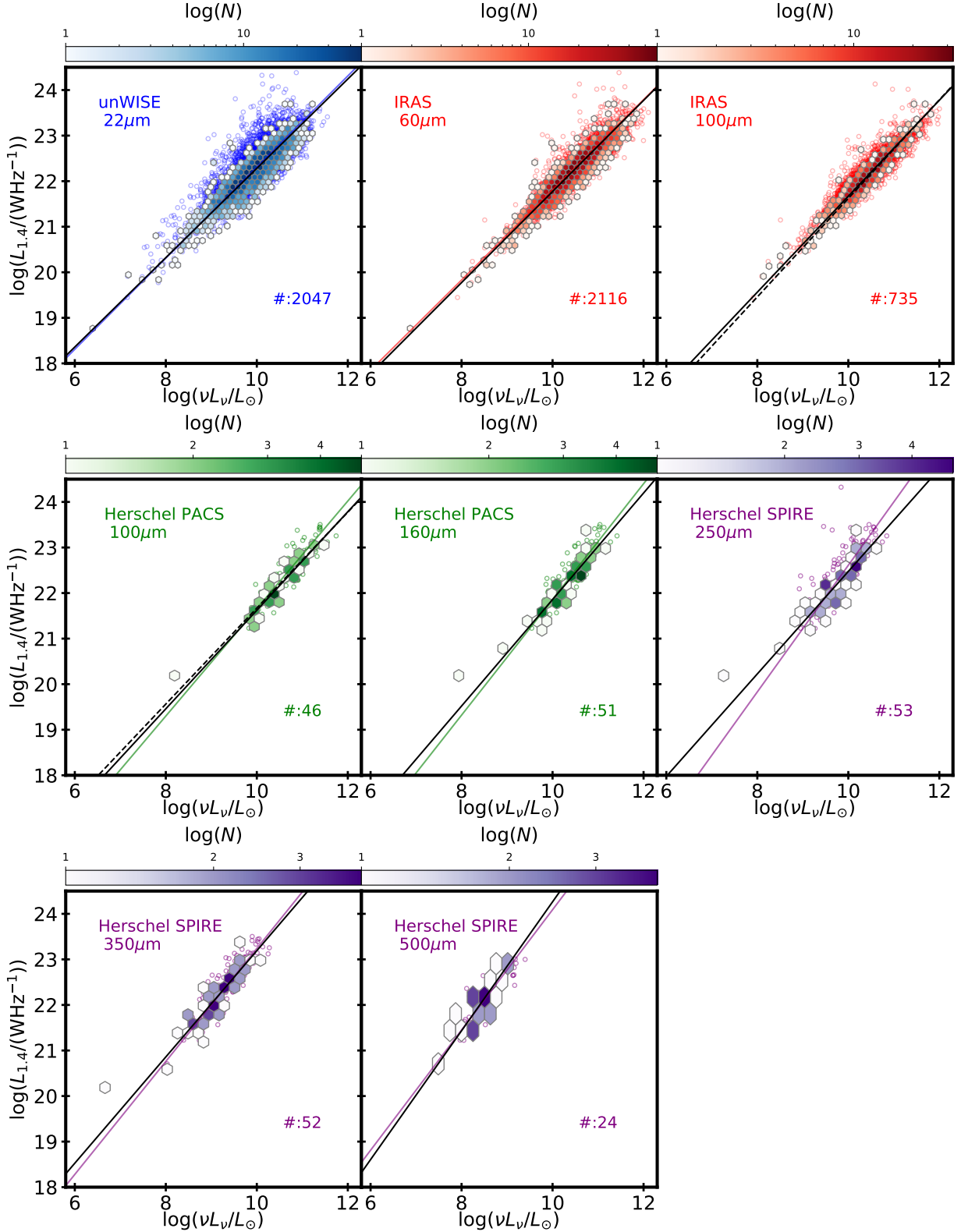


Figure 8. Radio luminosity at 1.4 GHz as a function of different monochromatic IR luminosities. Coloured hexagonal bins (empty circles) represent sources in the depth-matched SFG sample (full combined sample) with a >5 SNR detection in the IR band of the respective panel (see Table 4 for the selection criteria in each band). Black lines are the best-fit linear relations using depth-matched SFGs, while coloured lines are fits to the full combined sample. The dashed lines in the two $100\mu\text{m}$ panels show the best fit, respectively, to the measurements of the other $100\mu\text{m}$ data set. The number of sources used in the fit is specified in the bottom right quarter of each panel. Table 5 reports all best-fit slope and dispersion values for the depth-matched SFGs, and additionally also for the combined sample and the entire depth-matched sample.

colder dust components in the interstellar medium below and above $100\ \mu\text{m}$, respectively. At the same time, we see a decrease in best-fit intercepts (b_λ) which reaches a minimum around $100\text{--}160\ \mu\text{m}$, and then again rises up to $350\ \mu\text{m}$ (see lower panel of Fig. 9). Considering the broadly similar slope values at all wavelengths (within 20%), to zeroth order we expect the intercepts to generally reflect the changing SED amplitude at the respective wavelengths, with the addition that our choice to fit $L_{1.4}$ as a function of νL_ν mirrored this trend, and thus for visualization purposes we inverted the y-axis of in the bottom panel of Fig. 9.

With this physical picture in mind we empirically approximated the wavelength (λ) dependence of m_λ as:

$$m_\lambda = 0.08 \cdot \tanh \left[0.04 \cdot \left(\frac{\lambda}{\mu\text{m}} - 100 \right) \right] + 1.07, \quad (8)$$

and the λ vs. b_λ data with a grey body-like model:

$$b_\lambda = \log(4 \times 10^6) - \log \left(\frac{\lambda^{-1.9}}{(e^{175/\lambda} - 1)} \right). \quad (9)$$

The uncertainties on our best-fit monochromatic IRRC measurements are strongly dominated by sample sizes, and thus for optimizing the models that describe their λ dependence, we weighted each point equally. As a result, our trends should be considered tentative exploratory models rather than fully-realised fits. Furthermore, as seen in Fig. 9, the $500\ \mu\text{m}$ measurement is an outlier to the general behaviour of other bands. According to our simple physical interpretation outlined above, it should have a slope similar to the 250 and $350\ \mu\text{m}$ measurements, and correspondingly a smaller intercept than measured at $350\ \mu\text{m}$. We note, however, that this band is by far the most sensitive to the depth matching approach we employ due to the small numbers of detections involved. Indeed, a small change of ~ 0.1 dex in the flux density cut applied to the 1.4 GHz detections alters the best-fit m_{500} and b_{500} values such that they become consistent with the estimates in the other SPIRE bands, and such that the $500\ \mu\text{m}$ measurement would conform much better to our simple physical interpretation of the slope and intercept variations of monochromatic IRRCs.

Ideally, an analysis of the monochromatic IRRC parameters as a function of wavelength would be based on datasets of similar sizes, or even the same galaxies altogether at all wavelengths. Nevertheless, our Eqs 8 and 9 provide a tool to estimate the properties of monochromatic IRRCs across a wide IR wavelength range. In particular, this will be useful when studying high- z galaxies. At sub-millimetre wavelengths instruments such as the Atacama Large Millimeter Array or the James Clerk Maxwell Telescope sample the peak of IR SED of redshift of $2 - 3$ galaxies. Meanwhile sub- 1 GHz observations from the Giant Metrewave Radio Telescope, the Low-Frequency Array or MeerKAT are capable of detecting 1.4 GHz radio emission from the same sources. One could envisage even more combinations of monochromatic IR and radio observations across a wide redshift range, which will be compatible with each other through empirical formulae such as Eqs 8 and 9, and as a result, will help constraining the evolution of the IRRC already from suitably chosen single-band observations.

Finally, to facilitate comparison with other studies of

the monochromatic IRRC, we computed the more commonly used $q_\nu = \log(L_\nu/L_{1.4})$ values in each band. In Table 6 we report our median q_ν measurements derived in two different ways. On the one hand, we used our best-fit monochromatic IRRC models (parametrised by m_{HSFG} and b_{HSFG} from Table 5) and substituted the median $\log(L_\nu)$ value into Eq. 7 to calculate the expected median $\log(L_{1.4})$ and consequently median q_ν in each band (denoted as $\bar{q}_{\nu,\text{fit}}$ in Table 6). On the other hand, we also calculated individual q_ν values for galaxies shown in Fig. 8 and measured the median of their resulting distribution (\bar{q}_ν in Table 6). Due to the small sample size of the depth-matched H-ATLAS subsets, the statistical error on the median $\log(L_\nu)$ is quite large, leading to large uncertainties on the $\bar{q}_{\nu,\text{fit}}$ values for Herschel bands. Nevertheless, our $\bar{q}_{250} = 2.03 \pm 0.03$, calculated as the median of all q_{250} in the depth-matched $250\ \mu\text{m}$ sample matches well with the $\bar{q}_{250} = 2.01 \pm 0.04$ of Jarvis et al. (2010). The $q_{250} = 1.95 \pm 0.2$ measurement of Gürkan et al. (2018) is also compatible with our results. We note that their measured $250\ \mu\text{m}$ IRRC slope of 0.96 ± 0.01 is shallower than our 1.12 ± 0.12 , but still consistent within $\sim 1.5\ \sigma$. Meanwhile, Read et al. (2018) reported a higher $q_{250} = 2.30 \pm 0.04$ using the same catalogue as Gürkan et al. (2018), and the $q_{250} \approx 2.61$ of Smith et al. (2014) is even more offset from our median q_{250} . However, the latter study was based on a sample selected at $250\ \mu\text{m}$, and thus it is likely biased towards high q_{250} values. As this example shows, and for reasons explained in the following Sect. and in Sect. 4.3 and 5.1, we caution against comparing median IR-radio ratios without considering the IR and radio luminosity coverage of particular samples.

4.2 The total infrared-radio correlation

A general parametrization of the IRRC is possible via fitting the equation

$$\log \left(\frac{L_{1.4}}{\text{W Hz}^{-1}} \right) = M \cdot \log \left(\frac{L_{\text{TIR}}}{L_\odot} \right) + B. \quad (10)$$

With the use of same BCES fitting methodology as employed for characterising the monochromatic IRRCs (Sect. 4.1), and considering the 2,047 SFGs with reliable SED models in our depth-matched sample we obtain $M = 1.114 \pm 0.009$ (see left-hand panel of Fig. 10), consistent with the 1.10 ± 0.04 measurement of Bell (2003). This greater than unity slope implies a non-linear IRRC, which has consequences for the most widely used metric of the IRRC, q_{TIR} , defined as the logarithmic ratio, of the total infrared and 1.4 GHz radio ($L_{1.4}$) luminosities:

$$q_{\text{TIR}} \equiv \log \left(\frac{L_{\text{TIR}}}{3.75 \times 10^{12} \text{ W}} \right) - \log \left(\frac{L_{1.4}}{\text{W Hz}^{-1}} \right). \quad (11)$$

Median or mean q_{TIR} values are often used to characterize the IRRC properties of a sample of galaxies throughout the literature. Based on Eq. 11, constant q_{TIR} values lie alongside lines with slopes of unity in the $L_{\text{TIR}} - L_{1.4}$ parameter space. Our estimated slope of ~ 1.11 therefore suggests that the average q_{TIR} is decreasing towards higher luminosities, as seen in Fig. 10 and found by e.g. Jarvis et al. (2010); Morić et al. (2010); Ivison et al. (2010b) and Basu et al. (2015). Our best-fit model shown in red indeed connects

Table 5. Slope measurements (m_{all} , m_{dm} and m_{dmSFG}), intercepts (b_{all} , b_{dm} and b_{dmSFG}) and scatters (σ_{all} , σ_{dm} and σ_{dmSFG}) of each monochromatic IR – 1.4 GHz radio luminosity correlation, and the total IRRC for the combined sample, the full depth-matched sample and for the subset of depth-matched SFGs.

	$L_{22\mu\text{m}}$	$L_{60\mu\text{m}}$	$L_{100\mu\text{m,IRAS}}$	$L_{100\mu\text{m,H-ATLAS}}$	$L_{160\mu\text{m}}$	$L_{250\mu\text{m}}$	$L_{350\mu\text{m}}$	$L_{500\mu\text{m}}$	L_{TIR}
m_{all}	1.01 ± 0.01	0.99 ± 0.00	1.00 ± 0.06	1.18 ± 0.07	1.26 ± 0.08	1.36 ± 0.13	1.23 ± 0.08	1.30 ± 0.07	1.170 ± 0.009
b_{all}	12.30 ± 0.09	11.90 ± 0.05	11.63 ± 0.60	9.83 ± 0.70	9.32 ± 0.83	9.05 ± 1.31	10.98 ± 0.76	11.08 ± 0.61	9.7 ± 0.1
σ_{all}	0.22	0.14	0.17	0.15	0.15	0.23	0.17	0.15	0.18
m_{dm}	0.97 ± 0.01	1.00 ± 0.01	0.99 ± 0.07	1.08 ± 0.06	1.22 ± 0.08	1.19 ± 0.08	1.22 ± 0.08	1.61 ± 0.20	1.15 ± 0.01
b_{dm}	12.54 ± 0.09	11.73 ± 0.06	11.77 ± 0.75	10.83 ± 0.60	9.68 ± 0.82	10.61 ± 0.81	11.04 ± 0.74	8.54 ± 1.72	9.9 ± 0.1
σ_{dm}	0.17	0.17	0.24	0.12	0.14	0.15	0.16	0.15	0.14
m_{dmSFG}	0.98 ± 0.01	1.01 ± 0.01	1.06 ± 0.01	1.05 ± 0.09	1.17 ± 0.13	1.12 ± 0.12	1.21 ± 0.14	1.63 ± 0.21	1.114 ± 0.009
b_{dmSFG}	12.48 ± 0.09	11.69 ± 0.08	11.09 ± 0.12	11.11 ± 0.92	10.13 ± 1.33	11.27 ± 1.18	11.14 ± 1.30	8.31 ± 1.78	10.2 ± 0.1
σ_{dmSFG}	0.16	0.12	0.09	0.12	0.14	0.16	0.18	0.10	0.12

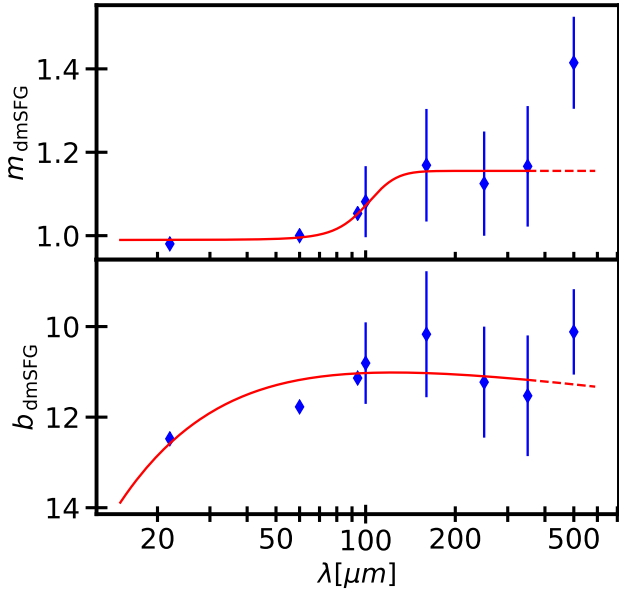


Figure 9. Measured monochromatic IRRC slopes and intercepts from Table 5/ Fig. 8 as a function of their wavelengths. Empirical approximations (red lines) give a calibration for any $\nu L_{\nu} - L_{1.4}$ relation between 22 and 500 μm . We plot the IRAS 100 μm data points at 94 μm to visually demonstrate the difference in uncertainties between IRRC parameters from Herschel and IRAS at 100 μm . Our choice to fit $L_{1.4}$ as a function of νL_{ν} mirrored the grey body-like trend in the bottom panel, and thus for visualization purposes we inverted the y-axis.

two dashed lines of constant q_{TIR} values with a difference of 0.6 dex across the ~ 4 dex luminosity range of our data. Therefore, median q_{TIR} values are dependent on the luminosity range of a given galaxy sample. In Sect. 5.1 we explore the implications of this in more detail.

Another effect on the q_{TIR} statistics demonstrated in Fig. 10 is related to the IR and radio sensitivity of a particular dataset, as mentioned in Sect. 3.3 and described in Sargent et al. (2010a). For example, the FIRST survey at $z = 0.01$ is typically sensitive to galaxies that lie above the black horizontal line in Fig. 10, while IRAS 100 μm is likely to detect sources to the right of the red vertical line. If we require a detection by both surveys, the resulting sample will likely miss several galaxies on the low q_{TIR} region above our best-fit model, and therefore have a median q_{TIR} biased

IR survey	$\log(\nu L_{\nu}/L_{\odot})$	$\bar{q}_{\nu,\text{fit}}$	\bar{q}_{ν}
WISE 22 μm	9.90 ± 0.01	1.2 ± 0.1	1.140 ± 0.006
IRAS 60 μm	10.39 ± 0.01	2.11 ± 0.08	2.098 ± 0.005
IRAS 100 μm	10.43 ± 0.02	2.4 ± 0.1	2.414 ± 0.006
H-ATLAS 100 μm	10.4 ± 0.1	2.5 ± 0.9	2.45 ± 0.04
H-ATLAS 160 μm	10.30 ± 0.09	2.4 ± 1.4	2.42 ± 0.04
H-ATLAS 250 μm	9.76 ± 0.09	2.0 ± 1.2	2.03 ± 0.04
H-ATLAS 350 μm	9.10 ± 0.09	1.6 ± 1.3	1.60 ± 0.04
H-ATLAS 500 μm	8.41 ± 0.09	1.2 ± 0.9	1.21 ± 0.06

Table 6. Median monochromatic luminosities ($\log(\nu L_{\nu}/L_{\odot})$), and median monochromatic IR-radio ratios calculated (a) using the best-fit relations (Eq 7) to the monochromatic IRRC in Fig. 8 ($\bar{q}_{\nu,\text{fit}}$) and (b) by taking the median of each q_{ν} distribution (\bar{q}_{ν}). Errors on (a) are computed by propagating the uncertainties on the best-fit slopes and intercepts (Table 5, final three rows) and the uncertainty on the median monochromatic luminosity, while for (b) we measured the standard error on the median of the q_{ν} distributions.

towards higher values. The magnitude of this bias is dependent on the mismatch between the depth of the IR and radio data. To obtain our depth-matched sample, we applied a 22 μm flux density cut of 19.5 mJy in order for a source to enter our IR-detected sample. This shifted the nominal L_{TIR} sensitivity of the unWISE catalog at $z = 0.01$ (shown as solid blue line) to a higher L_{TIR} (marked by the vertical dashed blue line) and matches the FIRST sensitivity well, in that the intersection of the lines of limiting L_{TIR} and $L_{1.4}$ lies almost on top of the best-fit total IRRC model. In Sect. 4.3 we further discuss the quantitative impact of relative IR and radio survey depths on the median q_{TIR} .

Regardless of these potential issues affecting q_{TIR} measurements, which also complicate comparisons between results from different datasets, they remain the basis of many widely used $L_{1.4} - \text{SFR}$ conversions (e.g. Yun et al. 2001; Murphy et al. 2011; Delhaize et al. 2017). Specifically, with a typical/representative q_{TIR} value one can estimate the SFR of a galaxy via

$$\text{SFR} \propto 10^{q_{\text{TIR}}} L_{1.4}. \quad (12)$$

under the assumption that L_{TIR} is a good proxy of the total galaxy SFR (see further discussion of this in Sect. 5.2). Thus the dependencies of q_{TIR} on various galaxy properties and redshift are crucial for improving the accuracy of these radio-based SFR estimates.

In our combined sample, we measure a median value of

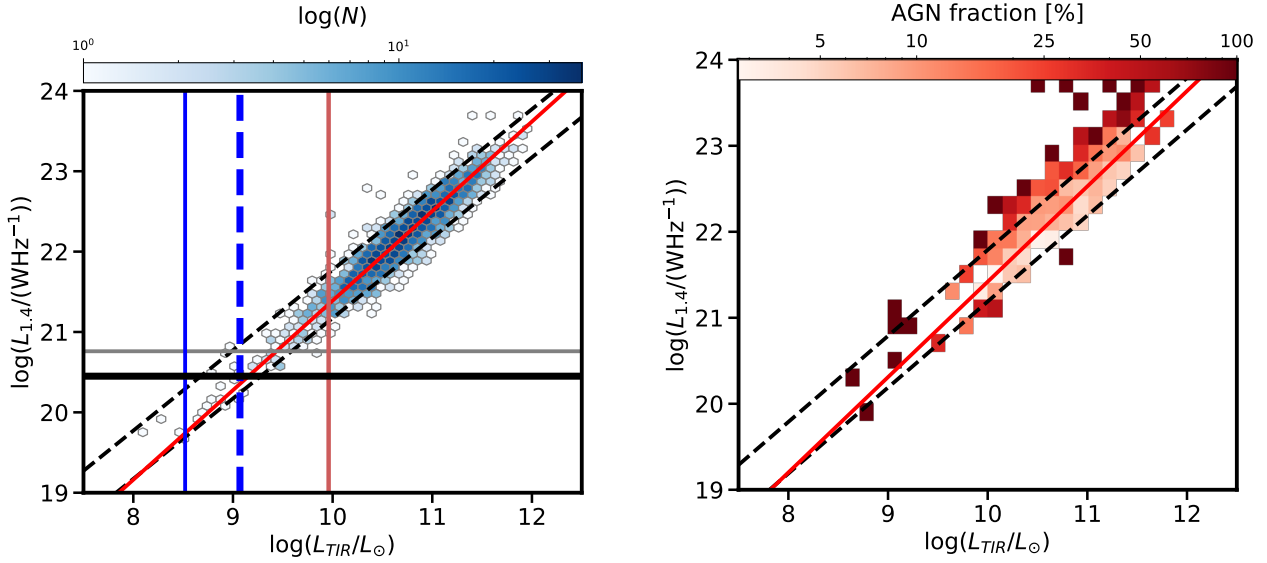


Figure 10. Left: 1.4 GHz luminosity as a function of the total infrared luminosity in the depth-matched SFG sample. The red line is our best-fit linear model, while dashed black lines are drawn at constant q_{TIR} values of $\bar{q}_{\text{TIR}} \pm 0.3$ dex. Table 5 lists the slope and scatter of the best-fit relation for the depth-matched SFGs, and additionally also for the combined sample and the entire depth-matched sample. The 2σ confidence band has a similar width as the best-fit line. Vertical and horizontal lines illustrate the depth of various surveys at $z = 0.01$: the blue solid line is the $\log(L_{\text{TIR}}/L_{\odot}) \sim 8.7$ limit of unWISE, the red vertical one is drawn at $\log(L_{\text{TIR}}/L_{\odot}) = 9.54$ for IRAS 100, and the horizontal black and grey lines are the FIRST and NVSS limits of $\log(L_{1.4}/(\text{WHz}^{-1})) = 20.35$ and 20.75 , respectively. The dashed blue line is the luminosity limit at $z = 0.01$ in our depth-matched sample.

Right: Fraction of AGN hosts on the $L_{\text{TIR}} - L_{1.4}$ plane in our depth matched sample after the removal of MIR AGN. The largest AGN fractions occur in the radio-loud regime (at high $L_{1.4}$, above the locus of the IRRC), and – due to the larger dispersion of their q_{TIR} distribution (Fig. 11) – also on the opposite, radio-quiet, side of the IRRC. The average AGN fraction of the combined sample is 9%.

$\bar{q}_{\text{TIR}} = 2.47 \pm 0.01$ with a scatter of 0.27 dex. In our depth-matched sample including AGN, SFGs, and unclassified objects, we find $\bar{q}_{\text{TIR}} = 2.51 \pm 0.01$, while the scatter is 0.22 dex. This 0.04 dex higher value is broadly consistent with predictions from Sect. 4.3 considering the ~ 0.65 dex sensitivity offset between unWISE $22 \mu\text{m}$ fluxes, our deepest and widest IR photometry, and the FIRST flux limit, as seen in Fig. 7. If we consider only SFGs in the depth-matched sample and thus remove most radio-loud objects, we find $\bar{q}_{\text{TIR}} = 2.54 \pm 0.01$ and a scatter of 0.17 dex¹⁶. This median measurement is ~ 0.1 dex lower than that by Bell (2003) and Yun et al. (2001). However, it is an excellent match to the average IR-radio ratio of $q_{\text{TIR}} = 2.52 \pm 0.03$ measured by Jarvis et al. (2010) using H-ATLAS and FIRST data and

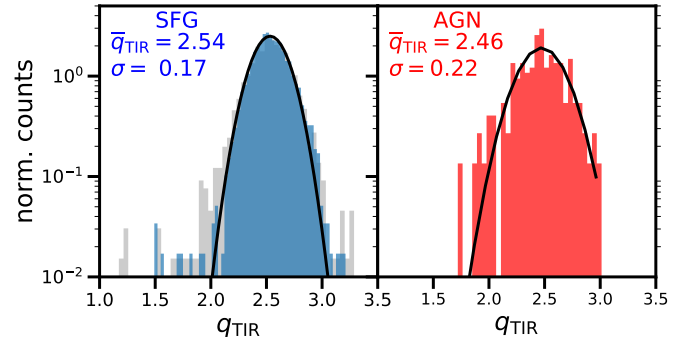


Figure 11. Distribution of q_{TIR} values for SFGs (left) and AGN (right) in the depth-matched, unbiased samples. In the left panel we show the q_{TIR} distribution of unclassified sources in the depth-matched sample as a grey histogram in the background. The ordinate axis is set to logarithmic scale. AGN show lower q_{TIR} on average, and larger scatter.

¹⁶ We note that the formal error on the median (estimated as $1.253\sigma/\sqrt{n}$, where σ is the standard deviation of the sample and n is the number of sources in the sample), is one order of magnitude smaller than the quoted value of 0.01 due to the large number of q_{TIR} measurements. However, systematic errors, such as the choice of SED template, do not permit a more precise determination of the median value. When the sample size is small enough for the formal error on the median to exceed this 0.01 dex threshold from systematics, we will quote a different error.

A (small) source of systematic error is the median q_{TIR} value assumed to match the sensitivity curves in Fig. 7 when defining our depth-matched sample. We have set our present flux limit assuming $\bar{q}_{\text{TIR}} = 2.64$ following Bell (2003). However, if we recalculate the flux cut according to our $\bar{q}_{\text{TIR}} = 2.54$, we only lower our \bar{q}_{TIR} in the SFG sample by 0.01 dex. This is consistent with the prediction of Fig. 12.

considering lower limits. The q_{TIR} distribution of SFGs in our depth-matched sample is shown in Fig. 11. A consequence of the non-linear IRRC is that the scatter of q_{TIR} in any given sample is systematically larger than the dispersion relative to the best-fit IRRC models of the same sample – in the case of the depth-matched SFGs these values are 0.17 dex and 0.12 dex, respectively.

Radio emission not related to the process of star-formation in AGN host galaxies leads, on average, to IR-radio ratios of AGN being lower than for pure SFG samples

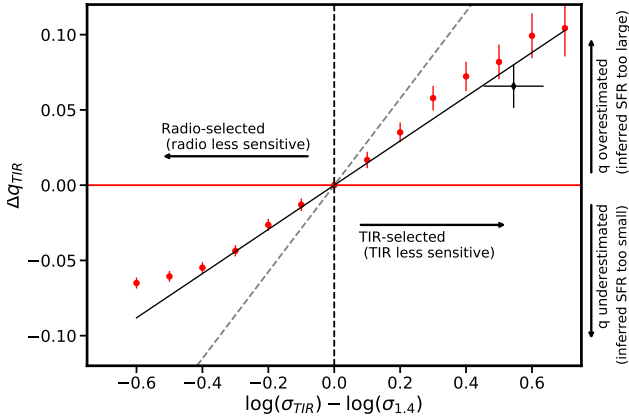


Figure 12. Median q_{TIR} difference to the median $q_{\text{TIR}} = 2.54$ of our depth-matched SFG sample as a function of offset between the sensitivity of TIR and radio data. The black point represents the measured offset of \bar{q}_{TIR} value for IRAS $60\mu\text{m} > 2\text{ Jy}$ sources (i.e. the selection used in Yun et al. 2001). It is broadly consistent with the bias predicted from our analysis. The dashed grey line is the $\Delta\bar{q}_{\text{TIR}}$ vs. $(\log(\bar{\sigma}_{\text{TIR}}) - \log(\bar{\sigma}_{1.4}))$ trend for a data set with a 40% larger scatter than our sample.

(see e.g. Ibar et al. 2008; Morić et al. 2010; Delhaize et al. 2017). Indeed, the AGN fraction on the $L_{\text{TIR}} - L_{1.4}$ plane, shown in the right panel of Fig. 10, is found to be higher in the radio excess region (above/to the left of the best-fit line), than on the main locus of the IRRC.

We measure a $\bar{q}_{\text{TIR}} = 2.46 \pm 0.02$ for AGNs in our depth-matched sample, with a scatter of 0.2 dex. This larger scatter compared to SFGs is also seen in Fig. 10, and is in qualitative agreement with the findings of e.g. Morić et al. (2010).

Composite sources have $\bar{q}_{\text{TIR}} = 2.54 \pm 0.01$ with a scatter of 0.2 dex. Since these sources likely harbour a complex mix of AGN, shock, and SF activity, we exclude them from further analysis and will concentrate on “pure” SFGs and AGN for the rest of this study.

4.3 The effect of flux limits on IRRC statistics

Sargent et al. (2010a) discuss how a mismatch in the sensitivities of IR and radio data can bias the median q_{TIR} value (\bar{q}_{TIR}) of a sample. In order to demonstrate the impact of such a difference in the flux limits, we created various flux-limited subsamples from our combined sample. Each of these subsamples was selected by applying different flux cuts either at $22\mu\text{m}$ or at 1.4 GHz. We quantified the mismatch between the IR and radio sensitivities of these samples, σ_{R} and $\sigma_{1.4}$ respectively, as the mean difference between the L_{TIR} curves calculated for their given $22\mu\text{m}$ and radio flux limits (see Fig. 7). We then measured \bar{q}_{TIR} in each of these subsamples and computed the q_{TIR} bias, $\Delta\bar{q}_{\text{TIR}}$, as the difference between the measured \bar{q}_{TIR} in the sample and $q_{\text{TIR}} = 2.54$, i.e. the median IR-radio ratio of our depth-matched SFG sample.

Fig. 12 shows $\Delta\bar{q}_{\text{TIR}}$ as a function of the logarithmic ratio of the IR and 1.4 GHz luminosity limits. At the origin of the figure is our depth-matched SFG sample with matched

IR and radio sensitivities. To the left of it, represented by negative $\log(\bar{\sigma}_{\text{TIR}}) - \log(\bar{\sigma}_{1.4})$ values, there are subsamples where we applied increasingly higher radio flux density cuts, and thus obtained $q_{\text{TIR}} < 2.54$ values. Conversely, in the right hand side of the figure, towards positive $\log(\bar{\sigma}_{\text{TIR}}) - \log(\bar{\sigma}_{1.4})$ values, we measure larger than 2.54 median IR-radio ratios. To this trend we fitted a linear model,

$$\Delta\bar{q}_{\text{TIR}} = m_{\text{bias}} \{ \log(\bar{\sigma}_{\text{TIR}}) - \log(\bar{\sigma}_{1.4}) \}, \quad (13)$$

with the slope m_{bias} , that has a best-fit value of 0.147 ± 0.007 . As discussed above, and in line with the analysis of Sargent et al. (2010a), we find that IR-limited (i.e. when radio data are more sensitive than IR) \bar{q}_{TIR} measurements are positively biased, whereas radio-limited samples lead to lower \bar{q}_{TIR} values. The magnitude of this bias is proportional to the mismatch between the radio and IR data. We note that this bias can be mitigated with either a selection based on a third, uncorrelated selection criterion, e.g. by studying a mass-selected sample, or techniques that allow probing IR and radio flux densities below their nominal limits, such as stacking or survival analysis, at least in the regime of not too strongly mismatched depths.

Our fit in Fig. 12 can in principle be used to quantitatively estimate and compare IRRC selection biases across different studies in the literature (Sargent et al. 2010a). However, in practice reconciling the median q_{TIR} values of different samples is not that straightforward. To see why, consider the following analytical formula for the difference between the average IR/radio ratio of IR- and radio-detected samples (Sargent et al. 2010a; see also analogous expressions in a variety of contexts in Kellermann 1964; Condon 1984; Francis 1993; Lauer et al. 2007):

$$\Delta q = \ln(10)(\beta - 1)\sigma_q^2. \quad (14)$$

Here β is the flux-dependent power-law index of the number counts (which, for the sake of simplicity, are assumed to have the same β in the IR and radio band when dealing with a pure SFG sample) and σ_q is the observed scatter of the IRRC. The power law indices of IR and radio number counts thus directly influence the value of $\Delta\bar{q}_{\text{TIR}}$, and since they are naturally a function of survey depth, Eq. 14 should not be viewed as producing a single, universal offset estimate, but instead has some dependence on luminosity. Similarly, changes in the observed scatter – to which both the intrinsic dispersion and measurement errors contribute – will change the slope of Eq. 13. To illustrate this, we show in Fig. 12 with a dashed grey line the $\Delta\bar{q}_{\text{TIR}}$ vs. $(\log(\bar{\sigma}_{\text{TIR}}) - \log(\bar{\sigma}_{1.4}))$ trend for a data set with a 40% larger scatter than our sample, resulting in a slope that is twice as steep. Nevertheless, if we bear in mind these different factors, Eq. 13 can still be used to reconcile apparently inconsistent results, and identify residual disagreement beyond the bias caused by this selection effect if present.

As an example, we consider the sample of Yun et al. (2001). Fig. 12 predicts that due to its unmatched selection criteria (for an illustration of this see Fig. 10), the median q_{FIR} of Yun et al. (2001) is biased high. In order to find evidence for this in our catalogue, we examined the \bar{q}_{FIR} values reported in Yun et al. (2001) ($\bar{q}_{\text{FIR, Yun}} = 2.34 \pm 0.01$) and the median q_{FIR} in a sample suitably depth-

matched between the IR and the radio at the level of the NVSS sensitivity curve in Fig. 7 ($\bar{q}_{\text{FIR, dm}} = 2.26 \pm 0.01$). For a fair comparison, we re-scaled the difference of these values, $\Delta q = 0.08$ dex, with the ratio of their scatters squared $(\sigma_{q, \text{dm, SFG}}/\sigma_{q, \text{Yun}})^2$ following Eq. 14 above. Substituting the dispersion of 0.26 dex reported in Yun et al. (2001), and our 0.23 dex scatter of q_{FIR} distribution (measured similar to Yun et al. (2001), i.e. not as the standard deviation of orthogonal distances which would result in a lower dispersion value) results in $\Delta q = 0.07$. The difference between average sensitivities in the Yun et al. (2001) catalogue was computed from the 2 Jy IRAS 60 μm and NVSS curves in Fig. 7, and its uncertainty was taken as the standard deviation of the differences in the plotted redshift range. The \bar{q}_{FIR} difference agrees within $\sim 1.5\sigma$ with our expectation based on Eq. 13. This suggests that the $\bar{q}_{\text{FIR}} = 2.34$ reported in Yun et al. (2001) is biased high and that the low- z \bar{q}_{FIR} value is instead nearer 2.26 in the relevant luminosity range.

Finally, we note that due to the different sample selection philosophy of Bell (2003) – who aim to maximise wavelength coverage from the far-ultraviolet to the radio, rather than basing sample selection on (a) tiered survey(s) – their data set does not lend itself to the same kind of systematic comparison we carried out above for the Yun et al. (2001) analysis. However, a more qualitative comparison is possible by considering the median luminosity $L_{\text{TIR}} = 10^{9.68} L_{\odot}$ of the Bell (2003) sample. At this luminosity, our IRRC best-fit parameters in Eq. 16 below translate to an IR-to-radio ratio of $q_{\text{TIR}} = 2.63$, which closely matches the median $q_{\text{TIR}} = 2.64 \pm 0.02$ of Bell (2003) and suggests that their sample – while situated in a lower luminosity regime than that of Yun et al. (2001), see Fig. 2 – is not subject to strong selection biases.

Fig. 12 also provides clues as to the potential issues with future studies seeking to investigate the IRRC using upcoming radio surveys if they are matched to already existing IR data. For example, from an IRRC perspective the targeted 1 μJy sensitivity of MIGHTEE (Jarvis et al. 2016) at $z < 0.2$ will be ~ 2.4 , 3.5 and 2.3 dex deeper than the unWISE 22 μm , IRAS 100 μm and Herschel 250 μm data we use in this paper, respectively. If the fitted trend in Fig. 12 is taken at face value, calculating \bar{q}_{TIR} based on a cross match between MeerKAT detections and these IR data without considering the different survey depths could result in \bar{q}_{TIR} estimates biased high by $\sim 0.25 - 0.4$ dex due to the significantly deeper radio observations. The 10 μJy detection limit of the EMU survey (Norris et al. 2011) with ASKAP will likely lead to a qualitatively similar bias, if not mitigated using an appropriate flux cut. With this newly arising large gap between IR and radio surveys, other calibration methods will become more important, e.g. the utilization of utilizing shorter wavelength or combined SFR tracers when exploiting deep radio data (see e.g. Hodge et al. 2008; Brown et al. 2017; Davies et al. 2017; Gürkan et al. 2018; Read et al. 2018; Duncan et al. 2020) as opposed to IR measurements. Arguably, for such low-luminosity sources, the IR emission may not do very well at capturing the bulk of the SFR anyway since typically their SFR-budget is dominated by the unobscured component.

5 DISCUSSION

5.1 The non-linearity of the IRRC

In this section we investigate the luminosity dependence of the IR/radio ratio, q_{TIR} . Fig. 13 presents q_{TIR} values for depth-matched SFGs and AGN host galaxies as a function of their radio and total IR luminosities. We fitted a linear model of the form of

$$q_{\text{TIR}} = s \log(L) + i, \quad (15)$$

where best-fit parameters s and i were found using the BCES method with orthogonal distance minimization. As the uncertainties on both IR and radio luminosities and q_{TIR} correlate strongly, we included their covariance in the BCES fits.

We find that q_{TIR} is only weakly dependent on L_{TIR} for both SFGs and AGN. The best-fit linear relation using depth-matched SFGs is

$$q_{\text{TIR, SFG}} = (-0.08 \pm 0.01) \cdot \log\left(\frac{L_{\text{TIR}}}{W}\right) + (3.4 \pm 0.1), \quad (16)$$

while for depth-matched AGN

$$q_{\text{TIR, AGN}} = (-0.14 \pm 0.05) \cdot \log\left(\frac{L_{\text{TIR}}}{W}\right) + (4 \pm 0.5). \quad (17)$$

Conversely, radio-bright, $\log(L_{1.4}/(W\text{Hz}^{-1})) \geq 22.5$, sources tend to have lower than average q_{TIR} values, while galaxies that are faint in radio have higher q_{TIR} . The best-fit relation between q_{TIR} and $\log(L_{1.4})$ in the depth-matched SFG sample is

$$q_{\text{TIR, SFG}} = (-0.177 \pm 0.009) \cdot \log\left(\frac{L_{1.4}}{W\text{Hz}^{-1}}\right) + (6.5 \pm 0.2), \quad (18)$$

while in depth-matched AGN sample it is

$$q_{\text{TIR, AGN}} = (-0.27 \pm 0.03) \cdot \log\left(\frac{L_{1.4}}{W\text{Hz}^{-1}}\right) + (8.4 \pm 0.7). \quad (19)$$

The fact that q_{TIR} appears to vary more with radio luminosity than with IR luminosity is consistent with the findings of Morić et al. (2010), Jarvis et al. (2010) and Ivison et al. (2010b). We note that even though AGN hosts and SFGs show qualitatively a similar behaviour, AGN have systematically steeper relations. This supports the scenario that the radio emission in AGN arises from different processes than for SFGs.

In Figs. 14 and 15 we present the q_{TIR} distributions of SFGs in the radio and IR luminosity bins defined in the 2nd row of Fig. 13. Fig. 16 shows the same information for AGN hosts (with bins defined as in the lower row of Fig. 13). These figures demonstrate a consistent Gaussian q_{TIR} profile around the fitted lines in the entire ~ 4 dex luminosity range. We observe a slightly decreasing scatter for SFGs both with increasing L_{TIR} and $L_{1.4}$. This is seemingly at odds with the findings of e.g. Yun et al. (2001), who measure a higher scatter at high radio and IR luminosities. However,

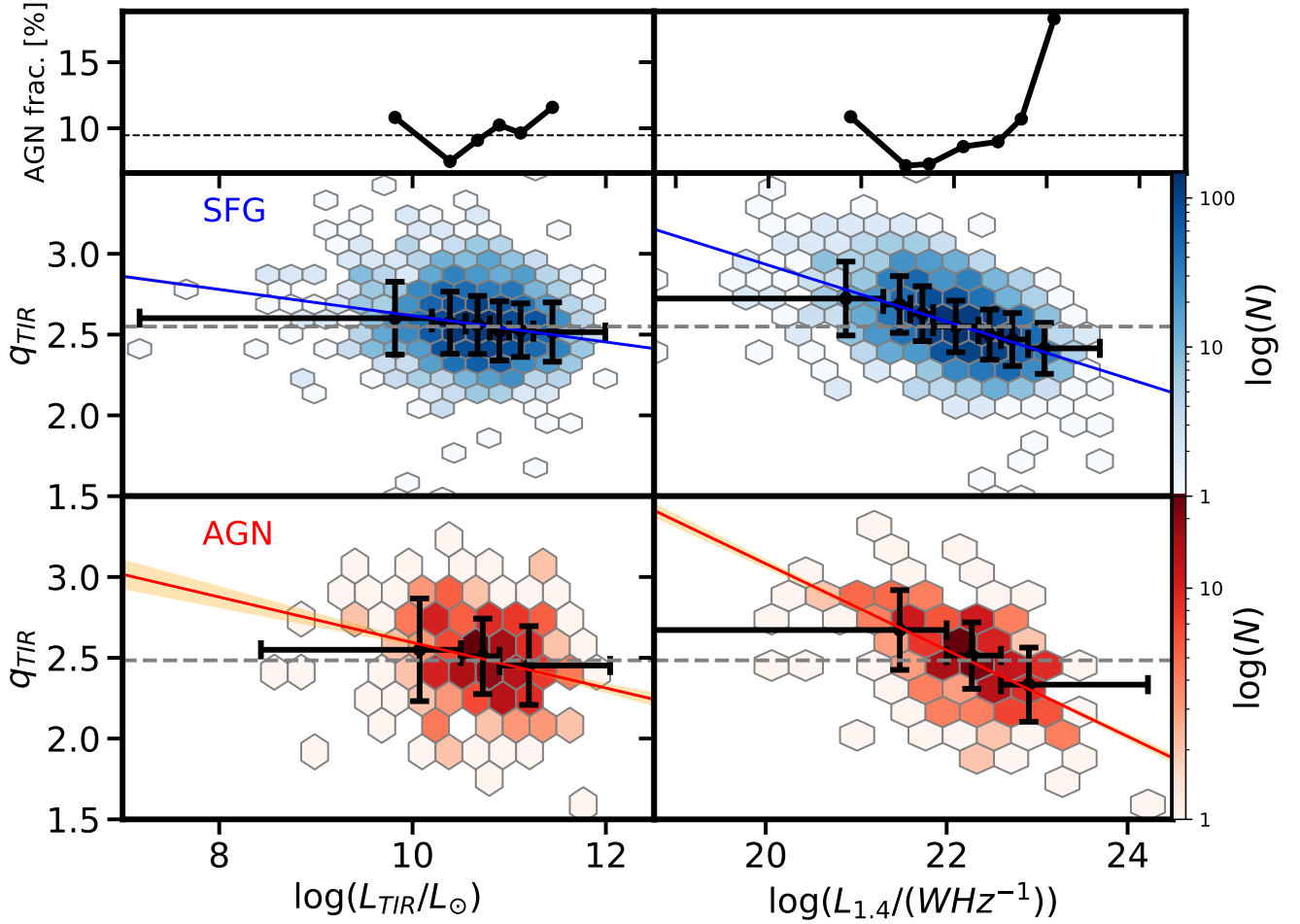


Figure 13. Median q_{TIR} of SFGs (middle) and AGN (bottom) as a function of IR (left) and radio (right) luminosity in the depth-matched sample. Black points are median q_{TIR} values of luminosity bins, vertical and horizontal error bars represent the measured scatter of these binned q_{TIR} distributions (for details see Figs. 14, 15 and 16) and the range of each luminosity bin, respectively. The dashed horizontal grey lines in these panels highlight the median q_{TIR} value found in the corresponding galaxy sample. Lines and shaded regions are the best fit and the 2σ uncertainty bands from our BCES fits, respectively. The upper row shows the variation of the AGN fraction with L_{TIR} and $L_{1.4}$. Fine dashed lines are drawn at the $\sim 9\%$ average AGN fraction of the depth-matched sample.

the sharply increasing AGN fraction towards this luminosity range both at IR and radio wavelengths (see Fig. 13 and 10) suggest that studies that do not separate these populations may find an artificially increased IRRC scatter especially at high luminosities, due to the on-average lower values and higher spread of AGN IR/radio ratios. Conversely, the increasing scatter towards low radio and IR luminosities may, at least to some extent, be caused by the expected breakdown of the IRRC due to UV and optical photons not being fully reprocessed by dust (see e.g. Bell 2003; Lacki et al. 2010).

As Morić et al. (2010) discussed, in a given sample, where q_{TIR} has no L_{TIR} dependence, and a non-zero dispersion, a declining trend with $L_{1.4}$ must be present due to the definition of q_{TIR} (for details on why this is the case see e.g. Condon 1984; Morić et al. 2010). To test whether the measured slope of the $q_{\text{TIR}} - L_{1.4}$ relation can be fully explained by this mathematical interdependence, Morić et al. (2010) generated mock data by sampling the observed q_{TIR} distribution

and calculated $L_{1.4}$ by combining these randomized q_{TIR} values with the real L_{TIR} measurements. Finally the bootstrapped $q_{\text{TIR}} - L_{1.4}$ relation was fitted. Indeed, having carried out this exercise using our q_{TIR} distribution in Fig. 11 we find a consistent $q_{\text{TIR}} - L_{1.4}$ trend with our data, with a slope of -0.14 ± 0.01 .

The luminosity dependence of average q_{TIR} values is linked to the non-linearity of the IRRC. By definition (see Eq. 11), lines of constant, luminosity-independent q_{TIR} values have slopes of unity in the $L_{\text{TIR}} - L_{1.4}$ plane (see Fig. 10). However, if, as we see in our data, at low radio luminosities the average IR-radio ratio is high, while at high radio luminosities q_{TIR} tends to be low, a fit across the whole range has to deviate from a slope of unity to connect these regions. Thus, rather than adopting a single, constant q_{TIR} value, the dependence on 1.4 GHz luminosity of $q_{\text{TIR}}(L_{1.4})$ should be incorporated into radio continuum based SFR estimates. Substituting Eq. 15 into Eq. 12 results in

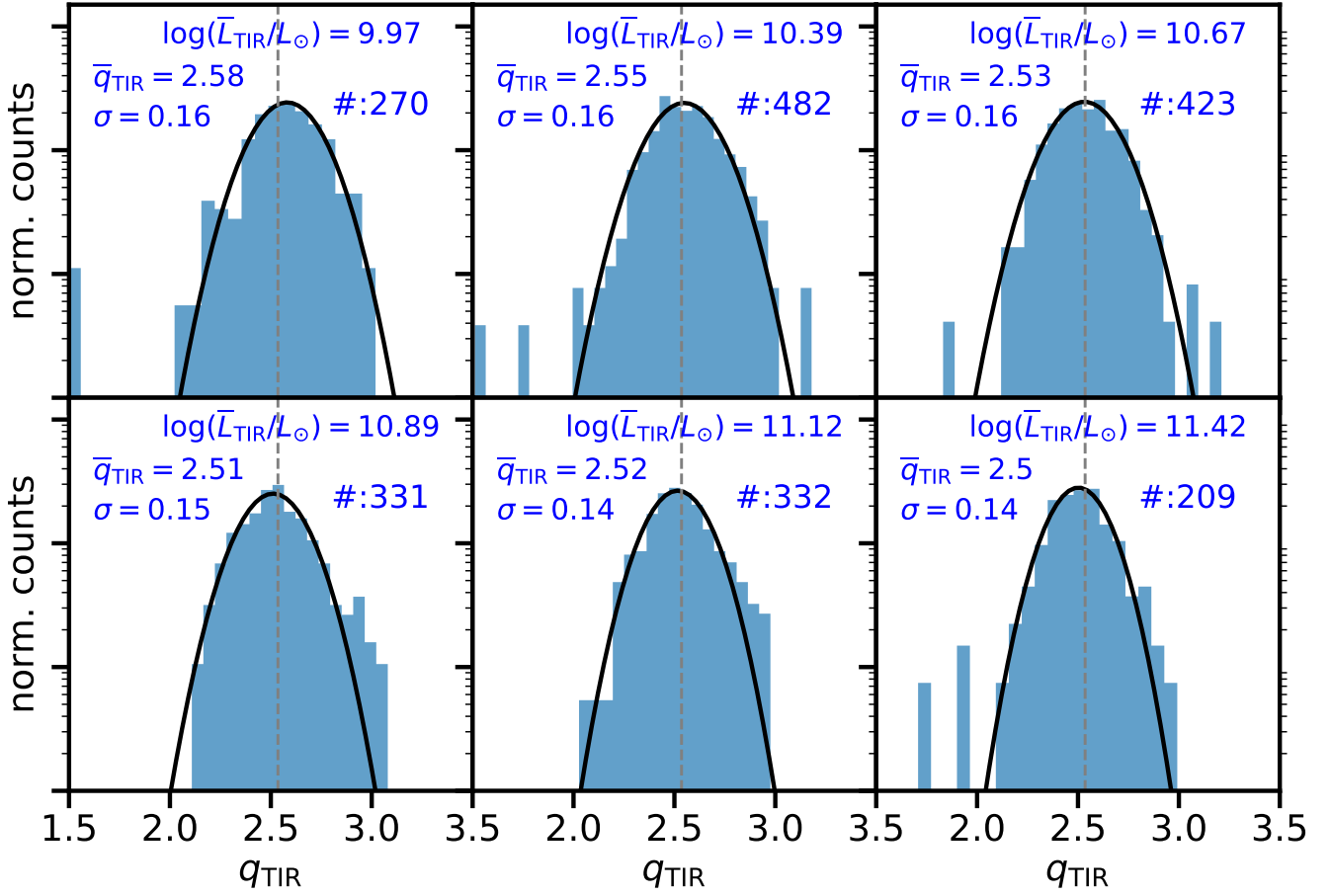


Figure 14. Distribution of q_{TIR} values for SFGs in the depth-matched sample in bins of increasing L_{TIR} (from the upper left to the lower right corner). Median IR luminosity values (\bar{L}_{TIR}), fitted median q_{TIR} (\bar{q}_{TIR}), scatter (σ) and number of sources in a given bin (#) are displayed in each panel, while bin widths are represented as horizontal errorbars in Fig. 13. Vertical, dashed grey lines represent the \bar{q}_{TIR} value of the entire depth-matched SFG sample for reference. The black curves are best-fit Gaussians to each distribution. The ordinate axis is set to logarithmic scale.

$$\text{SFR} \propto 10^{s \log(L_{1.4}) + i} L_{1.4}, \quad (20)$$

or, in log-log space

$$\log(\text{SFR}) = (s + 1) \log(L_{1.4}) + [i + C] \quad (21)$$

and with units and our best-fit parameters substituted

$$\log\left(\frac{\text{SFR}}{M_{\odot} \text{yr}^{-1}}\right) = (0.823 \pm 0.009) \cdot \log\left(\frac{L_{1.4}}{\text{WHz}^{-1}}\right) - (17.5 \pm 0.2), \quad (22)$$

assuming a $L_{\text{TIR}} - \text{SFR}$ scaling factor¹⁷ of $10^{-10} M_{\odot} \text{yr}^{-1} L_{\odot}^{-1}$.

¹⁷ We adopted the $4.5 \cdot 10^{-44} M_{\odot} \text{yr}^{-1} \text{erg}^{-1} s$ of Kennicutt 1998 (found in their Eq 4) and multiplied it by 0.61 (see e.g. Madau & Dickinson 2014) to account for the difference between the Salpeter (1955) IMF assumed by Kennicutt (1998) and the Chabrier (2003) IMF we adopt.

Since Eq. 20 assumes $\log(\text{SFR}) \propto \log(L_{\text{TIR}})$, we can see that its slope is related to the slope of the IRRC, where $\log(\text{SFR})$ is essentially re-scaled to $\log(L_{\text{TIR}})$. A comparison of Eq. 20 and Eq. 7 shows that the IRRC slope should therefore be roughly the inverse of the slope found for the $\log(\text{SFR}) - \log(L_{1.4})$ calibration. Indeed the latter is $(s + 1) \approx 0.85$, while the IRRC slope is 1.11 (reported in Table 5). Finally we note that as opposed to calibrations assuming a constant q_{TIR} value our $\log(\text{SFR}) - \log(L_{1.4})$ relation, and by extension the IRRC, is a power law. This contradicts the previously suggested conspiracy of $L_{1.4}$ and L_{TIR} to equally underestimate SFR in low luminosity galaxies (Bell 2003; Lacki et al. 2010).

5.2 1.4 GHz radio emission as a star-formation rate tracer

We tested the validity of an $L_{1.4}$ -dependent SFR calibration by comparing our recipe to radio-independent SFR estimates from the GALEX-SDSS-WISE Legacy Catalog (GSWLC; Salim et al. 2016). GSWLC SFRs were obtained via UV/optical SED fitting with CIGALE (Noll et al. 2009),

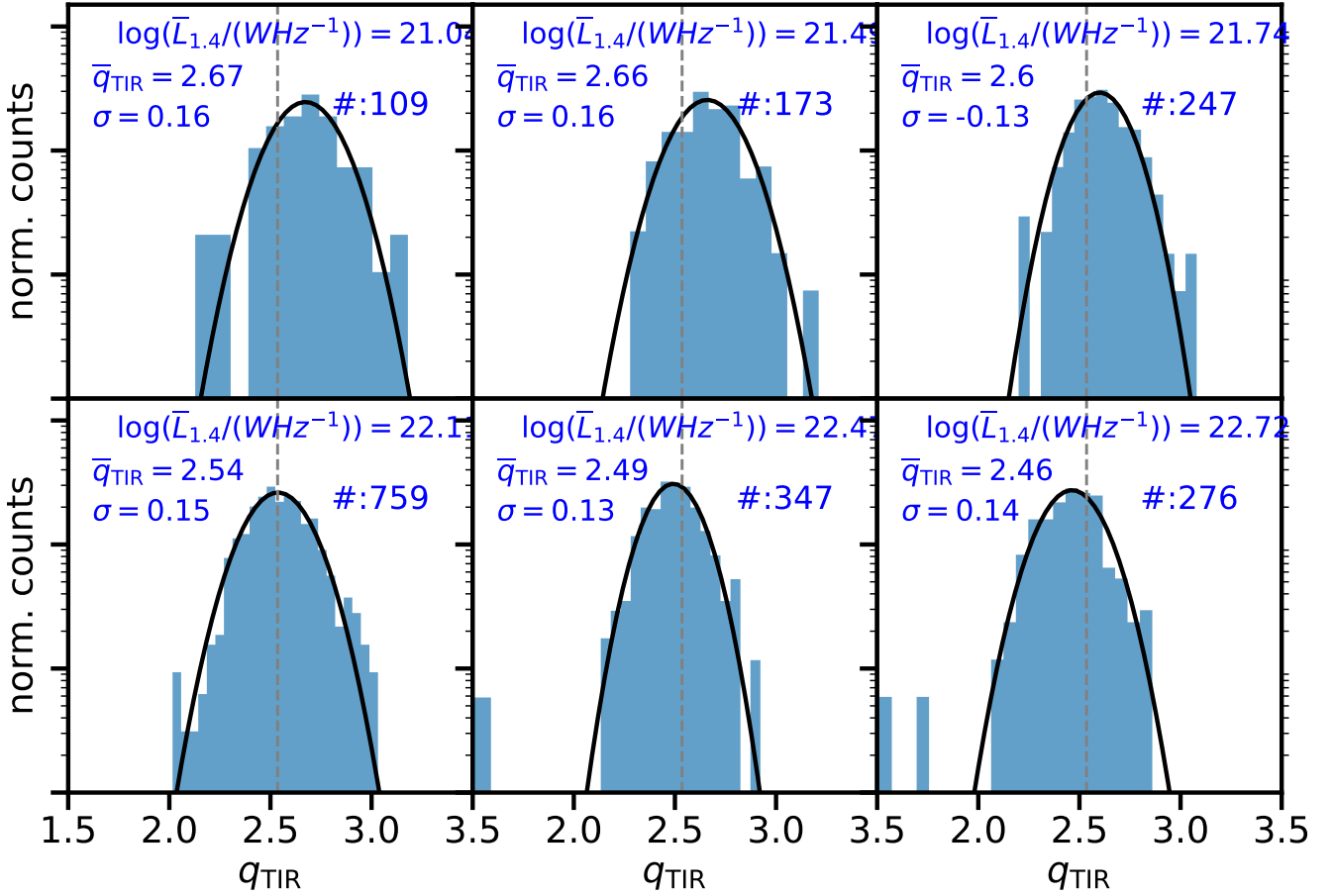


Figure 15. Same as Fig. 14, but with $L_{1.4}$.

independently of both TIR or radio luminosity. 1,740 of our depth-matched SFGs have SFR estimates in the GSWLC catalogue. In Fig. 17 we plot these against our $L_{1.4}$ measurements, and add several commonly used radio based SFR recipes (after conversion to a Chabrier 2003 IMF where necessary, see e.g. Madau & Dickinson (2014) for conversion factors). They have either been calibrated through the IRRC at low (Yun et al. 2001; Bell 2003, however, the latter applies a correction at $L_{1.4} < 6.4 \cdot 10^{21} \text{ WHz}^{-1}$) or high redshift (Delhaize et al. 2017, we use their Eq. 4 with the median redshift, $z=0.04$, of the galaxies shown in Fig. 17), or calibrated against non-IR tracers (Brown et al. 2017; Davies et al. 2017)¹⁸. We also show our $L_{1.4}$ -dependent IRRC SFR calibration (Eq. 22). The lower panel of Fig. 17 shows the mean offsets of SFRs in the GSWLC catalogue and SFRs estimated using the aforementioned $L_{1.4}$ – SFR calibrations.

In the $\log(L_{1.4}/\text{WHz}^{-1}) < 21.5$ regime our conversion is consistent with the Bell (2003), Brown et al. (2017) and Davies et al. (2017) formulae as well as the SED-derived SFRs at the $\sim 10\%$ level. On the other hand, the Yun et al. (2001) recipe predicts $\sim 25\%$ lower SFRs compared to the reference SFRs. Meanwhile, as a result of their

$L_{1.4}$ -independent q_{TIR} values, Yun et al. (2001) and Bell (2003) predict systematically higher SFR values in the range $\log(L_{1.4}/\text{WHz}^{-1}) > 22$, reaching an ~ 0.2 dex excess at $\log(L_{1.4}/\text{WHz}^{-1}) \approx 23$. In comparison, our conversion is consistent within $\sim 15\%$ with the GSWLC values, while Brown et al. (2017) and Davies et al. (2017) stay below 5 – 10%. Using the Delhaize et al. (2017) calibration on the other hand would yield 0.2 – 0.4 dex higher SFR estimate across the full luminosity range in this low- z sample, due to their median q_{TIR} being ~ 0.2 dex higher at $z \sim 0.04$ relative to the measurements in Yun et al. (2001) and Bell (2003), and ~ 0.3 dex higher than our $\bar{q}_{\text{TIR}} = 2.54$.

We note that it is typical to find 0.1-0.2 dex systematic offsets between various SFR estimates even at low redshift, implying that the normalisation of the offset curves in the lower panel of Fig. 17 is subject to a certain degree of systematic uncertainty. The key observation in this comparison is thus that the offset between SFRs calculated via our recipe in Eq. 22 and those from Salim et al. (2016) is essentially independent of radio luminosity, similar to Brown et al. (2017) and Davies et al. (2017), but unlike other IRRC-based methods, e.g. Yun et al. (2001); Bell (2003) and Delhaize et al. (2017). In conclusion, if we assume the SFRs based on the UV/optical SED fitting in Salim et al. (2016) are a robust benchmark (see their Sects. 7 and 8 for a com-

¹⁸ We used Eq 3 from Davies et al. (2017), i.e. 1.4 GHz radio luminosity calibrated against SED-fit derived SFRs.

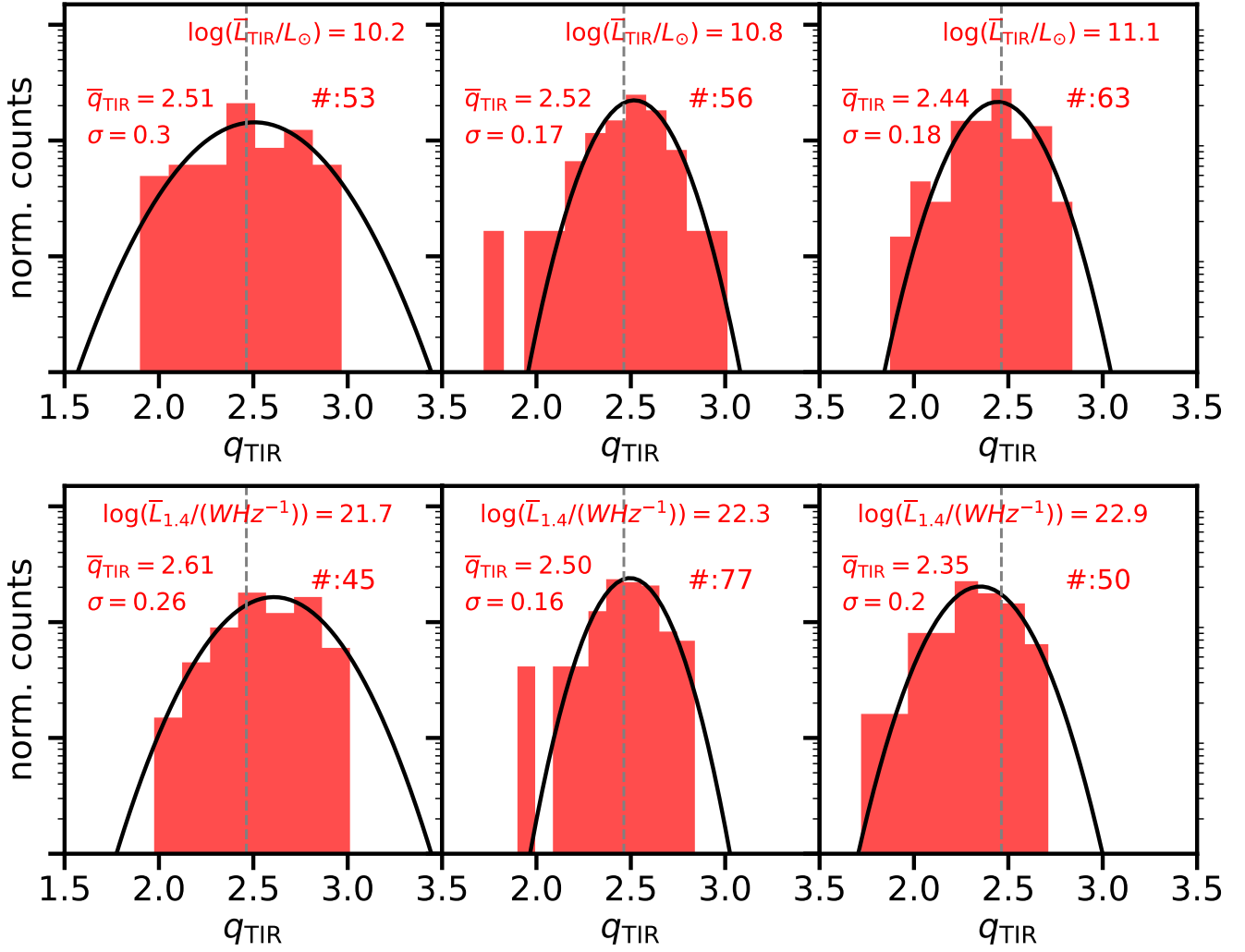


Figure 16. Same as Fig. 14 and 15, but for AGN host galaxies. The upper (lower) row shows q_{TIR} distributions for AGN hosts at different L_{TIR} ($L_{1.4}$), respectively.

parison to other widely used SFR measurement techniques and catalogues, respectively), IRRC-based $L_{1.4}$ – SFR conversions with a constant q_{TIR} that is independent of $L_{1.4}$ systematically underestimate SFRs in low-luminosity sources and overestimate them at high luminosities. The Bell (2003) recipe sought to resolve this issue by providing a modified prescription below $\log(L_{1.4}/W\text{Hz}^{-1}) \approx 21.8$, but remains less accurate than more recent calibrations at higher luminosities. Having dropped the assumption of a fixed q_{TIR} , our calibration – which is more akin to the approach of Hodge et al. (2008) – achieves a significantly better agreement with studies that do not solely use IR emission to infer SFRs.

It is important to note that, while our calibration performs well in the luminosity regime we probe ($\text{SFR} \gtrsim 0.5 M_{\odot}/\text{yr}$), a purely IRRC-based approach to calibrating SFRs becomes less and less tenable as one pushes to systems with lower mass and luminosity, where a much larger fraction of the star formation activity is not obscured by dust. Both these astrophysical reasons, as well as the pragmatic desire not to discard large numbers of faint sources due to dissimilar IR and radio survey depths (see discussion at the end of Sect.

4.3), imply that SFR measurements from multi-wavelength photometry or nebular emission lines will play an important role for the calibration of radio SFRs in deep radio surveys with SKA and its precursors (see Hodge et al. 2008; Davies et al. 2017; Brown et al. 2017; Gürkan et al. 2018; Duncan et al. 2020, for examples of such studies that have already pursued this approach using current radio data). Nevertheless, we still expect that there will continue to be applications where a purely IRRC-based calibration, and in particular a depth-matching approach as we discuss in this paper, remain useful. For example, this could be the case where a study focuses on a measurement of the scatter of the IRRC (which is not easily recoverable by stacking) in the high-luminosity regime in order to learn about the underlying physical processes, or when dealing with the rare population of highly dust-obscured starbursts, of which larger numbers will be picked up out to higher redshifts thanks to the higher survey speeds of the new generation of radio telescope arrays.

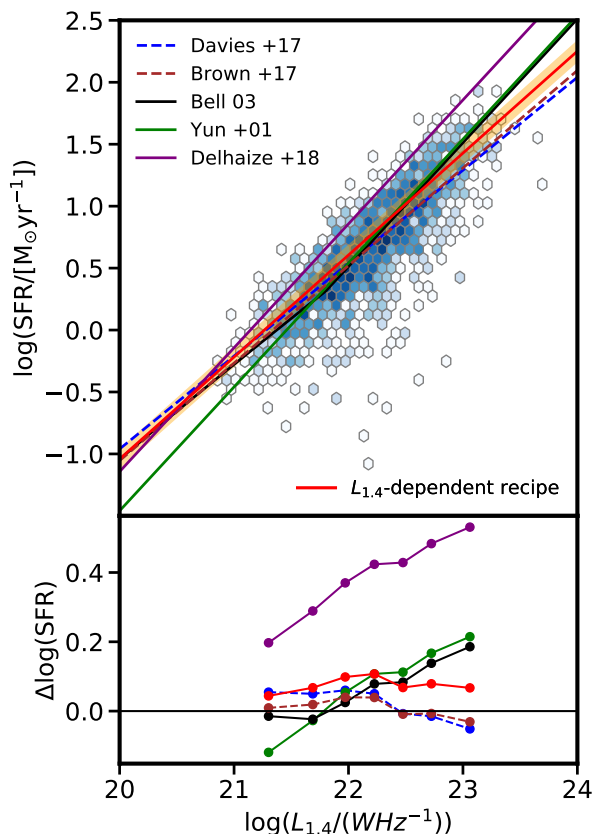


Figure 17. Top: The star formation rate – $L_{1.4}$ correlation using our depth-matched SFG sample. SFR values are taken from the GSWLC catalogue of Salim et al. (2016). Various models are presented as well as our best-fit to the data, and the SFR-conversion using our $L_{1.4}$ -dependent IR-radio ratio. The 1σ significance band around the latter was calculated from the correlated uncertainties on our $q_{\text{TIR}} - \log(L_{1.4})$ fit’s parameters. Bottom: Mean logarithmic ratio of the various SFR recipes and the reference SFR estimate of Salim et al. (2016) in bins of $L_{1.4}$.

5.3 The redshift dependence of the infrared-radio correlation

In the past decade it has been intensely debated whether the IRRC – and hence the relation between radio luminosity and SFR – evolves with redshift. Statements on the (lack of) evolution of the IRRC have almost exclusively been based on measurements of representative q -values for galaxy populations across different redshifts. A number of recent studies (Ivison et al. 2010a; Magnelli et al. 2015; Calistro Rivera et al. 2017; Delhaize et al. 2017) have found evidence for a declining radio-IR ratio across cosmic time (but see also Garrett 2002; Appleton et al. 2004; Garn et al. 2009; Jarvis et al. 2010; Sargent et al. 2010a,b; Mao et al. 2011; Smith et al. 2014; Pannella et al. 2015). Similar rates of modest, but statistically significant evolution have been reported at different observed frequencies, e.g., $(1+z)^{-(0.19 \pm 0.01)}$ at 3 GHz in Delhaize et al. (2017), $(1+z)^{-(0.12 \pm 0.04)}$ at 1.4 GHz in Magnelli et al. (2015), and in low-frequency LOFAR (van Haarlem et al. 2013) data Calistro Rivera et al. (2017) measured a consistent redshift dependency of $(1+z)^{-(0.15 \pm 0.03)}$ for SFGs. Since we have access to the according data, in the follow-

ing we will discuss the evolutionary trend of Delhaize et al. (2017) in the COSMOS field in more detail.

Delhaize et al. (2017) found no physical explanation for the trend of decreasing IR-to-radio ratios. Moreover, as discussed in Sect. 5.2, the extrapolation from the COSMOS sample in particular overestimates the $z = 0$ IR-radio ratios most commonly cited in the literature (Yun et al. 2001; Bell 2003). Using our $q_{\text{TIR}} - L_{1.4}$ relation we revisited the results of Delhaize et al. (2017). We computed median $L_{1.4}$ values in equal number redshift bins in the Delhaize et al. (2017) SFG sample using single-sided survival analysis¹⁹, and with the best-fit Eq. 18 we predicted the expected median q_{TIR} in each bin. Fig. 18 shows the median q_{TIR} values reported by Delhaize et al. (2017), and our predicted values. The 1σ confidence interval was calculated from the correlated slope and intercept uncertainties of our best-fit $q_{\text{TIR}} - L_{1.4}$ line combined with uncertainties on the median $L_{1.4}$ values in each redshift bin. Our empirical model qualitatively recovers the observed declining $q_{\text{TIR}} - z$ trend, albeit with a ~ 0.1 dex lower normalization and slightly shallower slope. In particular, compared to the $(1+z)^{-(0.19 \pm 0.01)}$ of Delhaize et al. (2017), our predicted $q_{\text{TIR}} - z$ fit for this COSMOS sample follows a $(1+z)^{-(0.16 \pm 0.01)}$ curve. Nevertheless, our model is typically in $\sim 1.5\sigma$ agreement with their measurements. Furthermore, it is an even better match to their radio-excess cleaned sample (shown in their Fig. 16 in cyan), which follows a $(1+z)^{-(0.15 \pm 0.01)}$ redshift evolution.

The preceding calculations suggest that the apparent redshift evolution of q_{TIR} is primarily a selection effect, as proposed by e.g. Basu et al. (2015), who find a slope identical to ours (1.11 ± 0.04) by sampling the IRRC at different luminosities with galaxies drawn from several redshift slices in the range $z < 1.2$. On the one hand there is the *physical* effect that, in the early Universe, galaxy star formation activity – and hence radio luminosity – was higher, which via Eq. 18 implies a lower q_{TIR} for high- z populations. On the other hand, *observational* selection effects cause galaxies fainter than the detection limit both in the radio and IR bands not to enter high- z samples, again skewing the measurement of IR-to-radio ratios towards high-luminosity objects with lower q values. To a varying extent, this Malmquist bias is present regardless of the details of the analysis/selection method, e.g. when considering only detected sources, but also when including upper flux limits for undetected sources. The combination of both effects produces a redshift-dependent sampling of an underlying non-linear relation, leading to declining IR-to-radio ratio measurements at higher redshifts.

Consistent with this interpretation of the redshift evolution of the IRRC reported in recent literature, the bivariate analysis carried out by Delvecchio et al. (2020) revealed that IR-to-radio ratios depend mostly on stellar mass and much less on redshift. As the SFR and stellar mass of SFG are correlated (e.g. Whitaker et al. 2012; Schreiber et al.

¹⁹ For a given a set of data containing both limits and direct measurements, survival analysis estimates the cumulative distribution function (CDF) of the underlying distribution they were drawn from. If, as in our particular case, besides direct detections, only either upper or lower limits occur (i.e. the data are singly censored), the CDF can be constrained analytically with the Kaplan–Meier product limit estimator (Kaplan & Meier 1958).

2015; Leslie et al. 2020), the preferential sampling of higher-luminosity, high-mass galaxies in the early Universe will lead to a qualitatively similar behaviour as outlined above. We also find a good quantitative agreement with the work of Delvecchio et al. (2020). For $M_{\star} = 10^{10} M_{\odot}$, the characteristic mass scale in their bivariate fitting formula (see Eq. (6) in Delvecchio et al. 2020) and at $z \sim 0$ these authors predicts $q_{\text{TIR}} = 2.65$. Given normalisation and slope of $z \sim 0$ literature star-forming main sequence fits (e.g. Brinchmann et al. 2004; Renzini & Peng 2015), $M_{\star} = 10^{10} M_{\odot}$ translates to an SFR of $\sim 1 M_{\odot}/\text{yr}$ or $L_{\text{TIR}} \sim 10^{10} L_{\odot}$. Using our IRRC best-fit parameters in Eq. 18, this luminosity implies an IR-to-radio ratio of $q_{\text{TIR}} = 2.67$, in excellent agreement with Delvecchio et al. (2020). In this context, the high $z=0$ q_{TIR} values in COSMOS in Delhaize et al. (2017) are thus likely due to the small volume probed at low redshifts, which causes few bright galaxies to enter the sample of Delhaize et al. (2017). As a result, the average $L_{1.4}$ in low- z bins is lower than that of wider surveys, with the consequence that the non-linear IRRC is probed in a regime with a higher effective q_{TIR} (see Fig. 10). By the same logic, the varying slopes of different $q_{\text{TIR}} - z$ fits in previously mentioned works are potentially related to their different observed radio luminosity distributions at each redshift.

In conclusion, we suggest that the various proposed $q_{\text{TIR}} - z$ calibrations in the literature so far are in general accurate for the data sets they were derived from. For instance, Delhaize et al. (2017) suggested a redshift-dependent $L_{1.4} - \text{SFR}$ calibration based on their declining q_{TIR} measurements. Novak et al. (2017) then applied this calibration to infer the cosmic SFR density (SFRD) out to $z \sim 5$ from 3 GHz radio data in COSMOS, finding a SFRD evolution that is broadly consistent with previous measurements. This is due to the fact that the underlying q -measurements by-and-large produce the correct relation between radio synchrotron luminosity and SFR for the L^* population which contributes most to the SFRD. A more universally consistent approach, however, would be the use of Eq. 22 or a similar formula (e.g. Davies et al. 2017; Brown et al. 2017; Gürkan et al. 2018; Duncan et al. 2020), until physically motivated models fitting observational data emerge.

6 SUMMARY

In order to provide a new, comprehensive study of the IRRC, and ultimately improve the $L_{1.4} - \text{SFR}$ calibration (crucial for the upcoming new generation of radio surveys with MeerKAT, ASKAP, and eventually SKA and ngVLA), we assemble an SDSS-based catalogue of 9,645 IR- and radio-detected galaxies in the nearby ($z < 0.2$) Universe. Thanks to our large initial pool of galaxies, even with stringent selection criteria excluding AGN or low-quality measurements, we retain $\sim 2,400$ SFGs in our final sample. To improve on previous similar works, we (i) utilize more recent IR surveys to achieve a better IR wavelength coverage and measure more accurate IR luminosities via SED-fitting, (ii) add deeper FIRST radio data to our catalogue, (iii) select pure star forming galaxies, (iv) consider the bias in the median IR-radio ratio arising from non-matching IR and radio survey depths and (v) employ a fitting technique to model the

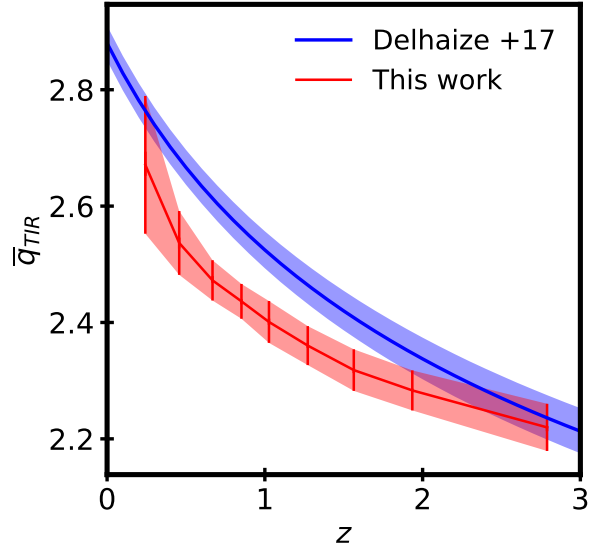


Figure 18. Redshift dependence of the IRRC. Blue symbols represent median \bar{q}_{TIR} values in the COSMOS field from Delhaize et al. (2017). Blue line with the shaded area is the best-fit from Delhaize et al. (2017) with 1σ confidence interval. Red line shows our predicted \bar{q}_{TIR} as a function of redshift based on the median $L_{1.4}$ luminosity of COSMOS SFGs in each z bin using Eq. 18. The red band shows the 1σ confidence interval of our prediction based on the uncertainties and covariance of our fit parameters and the median $L_{1.4}$ observed in COSMOS data.

IRRC that is shown to be more robust than the typical least square fit or bisector approaches.

With galaxy emission line ratios we separated pure star-forming galaxies from optically-selected AGN (Sect 3.2). Since our IR-SED fitting employs star-forming galaxy template SEDs, and some AGN can have substantial non-SF related IR emission, we also removed sources identified as AGN by their MIR colours. With the SFG sample we investigate selection effects biasing the median IR/radio ratio q_{TIR} , and which arise from poorly matched sensitivity levels in the radio and IR data. We were able to quantify the level of bias present in the widely referenced Yun et al. (2001) study and mitigated it through a “matched-depth” approach, i.e. by applying a $22 \mu\text{m}$ flux cut (see Sect 4.3). More generally, the details of any such bias corrections will depend on the flux/luminosity regime, but we demonstrate that they are important to consider. In particular, we expect that deep radio surveys such as MIGHTEE and EMU with MeerKAT and ASKAP (and ultimately SKA surveys), combined with existing IR data (e.g. from WISE, IRAS or Herschel) in the low- z Universe will obtain median q_{TIR} values biased high by $\sim 0.2 - 0.4$ dex, if the difference in sensitivities between IR and radio data is not taken into account.

In Sects 4.1 and 4.2 we characterise the IRRC properties of our depth-matched SFG sample using both monochromatic and total IR luminosities. We find that, in general, $22 - 100 \mu\text{m}$ monochromatic correlations have close to unity IRRC slopes, which becomes steeper towards longer wavelengths. The total IR-based correlation has a slope of 1.11 with a dispersion of 0.12 dex. This significant decrease in dispersion relative to previous literature is, in part, due to the difference in the dispersion measurement itself. In order

to carry out a fairer comparison, we also fitted our data considering L_{TIR} as the independent variable, and estimated the scatter as the standard deviation of the offsets from this line in the y -direction. This resulted in a ~ 0.18 dex scatter, suggesting a genuine improvement compared to the previously reported 0.26 dex scatter of Yun et al. (2001), regardless of the fitting approach.

Recipes for deriving SFRs from $L_{1.4}$ often involve the median IR-radio ratio which we also measure for the galaxies in our sample. In the total depth-matched sample we find $\bar{q}_{\text{TIR}} = 2.51$ with a scatter of 0.22 dex, for depth-matched SFGs we obtained $\bar{q}_{\text{TIR}} = 2.54 \pm 0.01$ and scatter of 0.19 dex, while depth-matched AGN have $\bar{q}_{\text{TIR}} = 2.46 \pm 0.02$ and a scatter of 0.27 dex. These scatters are systematically higher than those found for the IRRC itself, due to the non-unity slope of the correlation, which in turn is the result of q_{TIR} values (anti-)correlating with radio luminosity. Thus, instead of using a fixed q_{TIR} value, we propose an $L_{1.4}$ -dependent \bar{q}_{TIR} , and consequently, SFR calibration (Eq 22). In Sect 5.2 with IR- and radio-independent SFR estimates available for our SDSS sources, we confirm that such a recipe is compatible with other existing $L_{1.4}$ – SFR calibrations in the local Universe. More importantly, it goes a long way to empirically explain the apparent evolution of \bar{q}_{TIR} found by recent studies as the consequence of a selection effect whereby different parts of the non-linear IRRC are sampled depending on redshift and sample depth (Sect 5.3). Hence, it provides a robust $L_{1.4}$ – SFR recipe for both low and high-redshift 1.4 GHz radio observations.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for the thorough and insightful reports allowing us to improve this work. We thank Gábor Marton, Paolo Serra, Filippo Maccagni, Mpati Ramatsoku, Dane Kleiner, Scott Clay and Benoît Fournier for all the useful discussions. DCM acknowledges support from the Science and Technology Facilities Council (grant number ST/M503836/1). SL acknowledges funding from Deutsche Forschungsgemeinschaft (DFG) Grant BE 1837 / 13-1 r. MTS acknowledges support from a Royal Society Leverhulme Trust Senior Research Fellowship (LT150041). ES acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 694343). B.M. acknowledges support from the Collaborative Research Centre 956, sub-project A1, funded by the Deutsche Forschungsgemeinschaft (DFG) – project ID 184018867. JD acknowledges financial assistance from the South African Radio Astronomy Observatory (SARAO; www.ska.ac.za).

DATA AVAILABILITY

The data underlying this article are available in its online supplementary material. The same data are also available in Zenodo, at <https://zenodo.org/badge/latestdoi/344211798>.

REFERENCES

- Akras M. G., Bershadsky M. A., 1996, *ApJ*, 470, 706
 Alam S., et al., 2015, *ApJS*, 219, 12
 Andrae R., Schulze-Hartung T., Melchior P., 2010, arXiv e-prints, p. arXiv:1012.3754
 Appleton P. N., et al., 2004, *ApJS*, 154, 147
 Assef R. J., Stern D., Noirot G., Jun H. D., Cutri R. M., Eisenhardt P. R. M., 2018, *ApJS*, 234, 23
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
 Basu A., Wadadekar Y., Beelen A., Singh V., Archana K. N., Sirothia S., Ishwara-Chandra C. H., 2015, *ApJ*, 803, 51
 Becker R. H., White R. L., Helfand D. J., 1995, *ApJ*, 450, 559
 Bell E. F., 2003, *ApJ*, 586, 794
 Booth R. S., de Blok W. J. G., Jonas J. L., Fanaroff B., 2009, arXiv e-prints, p. arXiv:0910.2935
 Bourne N., et al., 2016, *MNRAS*, 462, 1714
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
 Brown M. J. I., et al., 2017, *ApJ*, 847, 136
 Calistro Rivera G., et al., 2017, *MNRAS*, 469, 3468
 Chabrier G., 2003, *PASP*, 115, 763
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
 Chary R., Elbaz D., 2001, *ApJ*, 556, 562
 Condon J. J., 1984, *ApJ*, 287, 461
 Condon J. J., 1992, *ARA&A*, 30, 575
 Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, *AJ*, 115, 1693
 Dale D. A., Helou G., 2002, *ApJ*, 576, 159
 Dale D. A., Helou G., Contursi A., Silbermann N. A., Kolhatkar S., 2001, *ApJ*, 549, 215
 Davies L. J. M., et al., 2017, *MNRAS*, 466, 2312
 DeBoer D. R., et al., 2009, *IEEE Proceedings*, 97, 1507
 Delhaize J., et al., 2017, *A&A*, 602, A4
 Delvecchio I., et al., 2020, arXiv e-prints, p. arXiv:2010.05510
 Driver S. P., et al., 2011, *MNRAS*, 413, 971
 Duncan K. J., Shivaee I., Shapley A. E., Reddy N. A., Mobasher B., Coil A. L., Kriek M., Siana B., 2020, *MNRAS*, 498, 3648
 Eales S., et al., 2010, *PASP*, 122, 499
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Francis P. J., 1993, *ApJ*, 407, 519
 Garn T., Green D. A., Riley J. M., Alexander P., 2009, *MNRAS*, 397, 1101
 Garrett M. A., 2002, *A&A*, 384, L19
 Griffin M. J., et al., 2010, *A&A*, 518, L3
 Gürkan G., et al., 2018, *MNRAS*, 475, 3010
 Helfand D. J., White R. L., Becker R. H., 2015, *ApJ*, 801, 26
 Helou G., Bica M. D., 1993, *ApJ*, 415, 93
 Helou G., Soifer B. T., Rowan-Robinson M., 1985, *ApJ*, 298, L7
 Hodge J. A., Becker R. H., White R. L., de Vries W. H., 2008, *AJ*, 136, 1097
 Hogg D. W., Bovy J., Lang D., 2010, arXiv e-prints, p. arXiv:1008.4686
 Ibar E., et al., 2008, *MNRAS*, 386, 953
 Ivison R. J., et al., 2010a, *MNRAS*, 402, 245
 Ivison R. J., et al., 2010b, *A&A*, 518, L31
 Jarrett T. H., Chester T., Cutri R., Schneider S. E., Huchra J. P., 2003, *AJ*, 125, 525
 Jarvis M. J., et al., 2010, *MNRAS*, 409, 92
 Jarvis M., et al., 2016, in *MeerKAT Science: On the Pathway to the SKA*. p. 6 (arXiv:1709.01901)
 Johnston S., et al., 2007, *Publ. Astron. Soc. Australia*, 24, 174
 Kaplan E. L., Meier P., 1958, *Journal of the American Statistical Association*, 53, 457
 Kauffmann G., et al., 2003, *MNRAS*, 346, 1055

Kellermann K. I., 1964, Publications of the Owens Valley Observatory, 1, 1

Kennicutt Jr. R. C., 1998, ARA&A, 36, 189

Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Kimball A. E., Ivezić Ž., 2008, AJ, 136, 684

Kimball A. E., Ivezić Ž., 2014, preprint, (arXiv:1401.1535)

Lacki B. C., Thompson T. A., Quataert E., 2010, ApJ, 717, 1

Lang D., Hogg D. W., Schlegel D. J., 2014, preprint, (arXiv:1410.7397)

Lang D., Hogg D. W., Schlegel D. J., 2016, The Astronomical Journal, 151, 36

Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, ApJ, 670, 249

Leslie S. K., Kewley L. J., Sanders D. B., Lee N., 2016, MNRAS, 455, L82

Leslie S. K., et al., 2020, ApJ, 899, 58

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Magnelli B., et al., 2015, A&A, 573, A45

Mao M. Y., Huynh M. T., Norris R. P., Dickinson M., Frayer D., Helou G., Monkiewicz J. A., 2011, ApJ, 731, 79

Marcillac D., Elbaz D., Chary R. R., Dickinson M., Galliano F., Morrison G., 2006, A&A, 451, 57

Marton G., et al., 2017, arXiv e-prints, p. arXiv:1705.05693

Molnár D. C., et al., 2018, MNRAS, 475, 827

Morić I., Smolčić V., Kimball A., Riechers D. A., Ivezić Ž., Scoville N., 2010, ApJ, 724, 779

Murphy E. J., et al., 2011, ApJ, 737, 67

Nemmen R. S., Georganopoulos M., Guiriec S., Meyer E. T., Gehrels N., Sambruna R. M., 2012, Science, 338, 1445

Neugebauer G., et al., 1984, ApJ, 278, L1

Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793

Norris R. P., et al., 2011, Publ. Astron. Soc. Australia, 28, 215

Novak M., et al., 2017, preprint, (arXiv:1703.09724)

Nyland K., et al., 2017, MNRAS, 464, 1029

Pannella M., et al., 2015, ApJ, 807, 141

Perley R. A., Chandler C. J., Butler B. J., Wrobel J. M., 2011, ApJ, 739, L1

Pilbratt G. L., et al., 2010, A&A, 518, L1

Poglitsch A., et al., 2010, A&A, 518, L2

Read S. C., et al., 2018, MNRAS, 480, 5625

Renzini A., Peng Y.-j., 2015, ApJ, 801, L29

Rigby E. E., et al., 2011, MNRAS, 415, 2336

Roychowdhury S., Chengalur J. N., 2012, MNRAS, 423, L127

Salim S., et al., 2016, ApJS, 227, 2

Salpeter E. E., 1955, ApJ, 121, 161

Sargent M. T., et al., 2010a, ApJS, 186, 341

Sargent M. T., et al., 2010b, ApJ, 714, L190

Schleicher D. R. G., Beck R., 2013, A&A, 556, A142

Schreiber C., et al., 2015, A&A, 575, A74

Smith D. J. B., et al., 2014, MNRAS, 445, 2232

Valiante E., et al., 2016, MNRAS, 462, 3146

Voelk H. J., 1989, A&A, 218, 67

Wang L., Rowan-Robinson M., Norberg P., Heinis S., Han J., 2014, MNRAS, 442, 2739

Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29

Wright E. L., et al., 2010, AJ, 140, 1868

Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803

de Jong T., Klein U., Wielebinski R., Wunderlich E., 1985, A&A, 147, L6

van Haarlem M. P., et al., 2013, A&A, 556, A2

van der Kruit P. C., 1971, A&A, 15, 110

van der Kruit P. C., 1973, A&A, 29, 263

APPENDIX A: DETERMINING SPSC AND PPSC MATCHING RADII AND ESTIMATING CONTAMINATION FROM SPURIOUS MATCHES

Since neither the SPSC nor the PPSC had already established optical counterparts with high accuracy positions (as opposed to, e.g., H-ATLAS), we determined band-by-band matching radii between Herschel and SDSS DR12 positions with the aim of minimising the spurious fraction while maximising the number of counterparts found.

To this end, we estimated the level of contamination as a function of matching radius in each band independently. First, we generated 15 mock IR catalogues for each band. Due to the non-contiguous coverage of the SPSC and PPSC, we used the positions of the real source IR catalogues as a starting point to simply create mocks that mimic the sky coverage of the real sources. The RA and Dec of each mock IR source was calculated by adding uniformly drawn random numbers between ± 18 arcsec to the real IR source positions. We then cross-matched the mock IR catalogues with the SDSS parent sample using TOPCAT positional cross match, taking the closest match out to a maximum matching radius of 15 arcsec. For these fake matches, it is more likely to find sources with larger separations, because a larger separation, or search radius, corresponds to a larger search circumference, consequently larger search area, and thus more random associations are possible. The distribution of optical-IR source separations for all 5 PSC bands is shown in Fig. A1. The blue histogram is the distribution of separations resulting from cross-matching our low- z parent catalogue with the real IR catalogues, and the orange histogram shows the result from one of the fake IR catalogue matches. For the spurious fraction estimations, we use the average distribution of all 15 fake catalogues. To calculate the contamination, we divide the number of fake sources (from the average of the fake catalogues) with the number of sources in the matched real catalogue that lie within our chosen search radius. The resulting curves are shown in the panels above each separation distribution. The search radii adopted for our final catalog cross-match corresponds to the radial-slice (rounded to the nearest half arcsecond) at which 50% of sources are spurious matches. The expected spurious fraction for each catalogue is $< 10\%$ except for Spire 350 and 500 micron that are 14% and 27%, respectively.

While the level of contamination among the 500 μm data is formally high, we note that spurious matches likely result in unphysical SED shapes, that are flagged by our method for filtering out poor SED models (see Sect. 3.1.1). Indeed, $\sim 16\%$ of all 500 μm -detected sources were identified as having unreliable model fits. Furthermore, $< 1\%$ of the 500 μm detections have only one other photometric point, and on average, these sources were observed in 6 bands, lowering the chances of a spurious data point significantly biasing their L_{TIR} estimate. Finally, due to 500 μm observations being relatively rare ($\sim 10\%$ of our combined sample), even if most sources with spurious 500 μm matches enter our depth-matched SFG sample (~ 100 sources), they are highly unlikely to distort our IIRC statistics in any meaningful way.

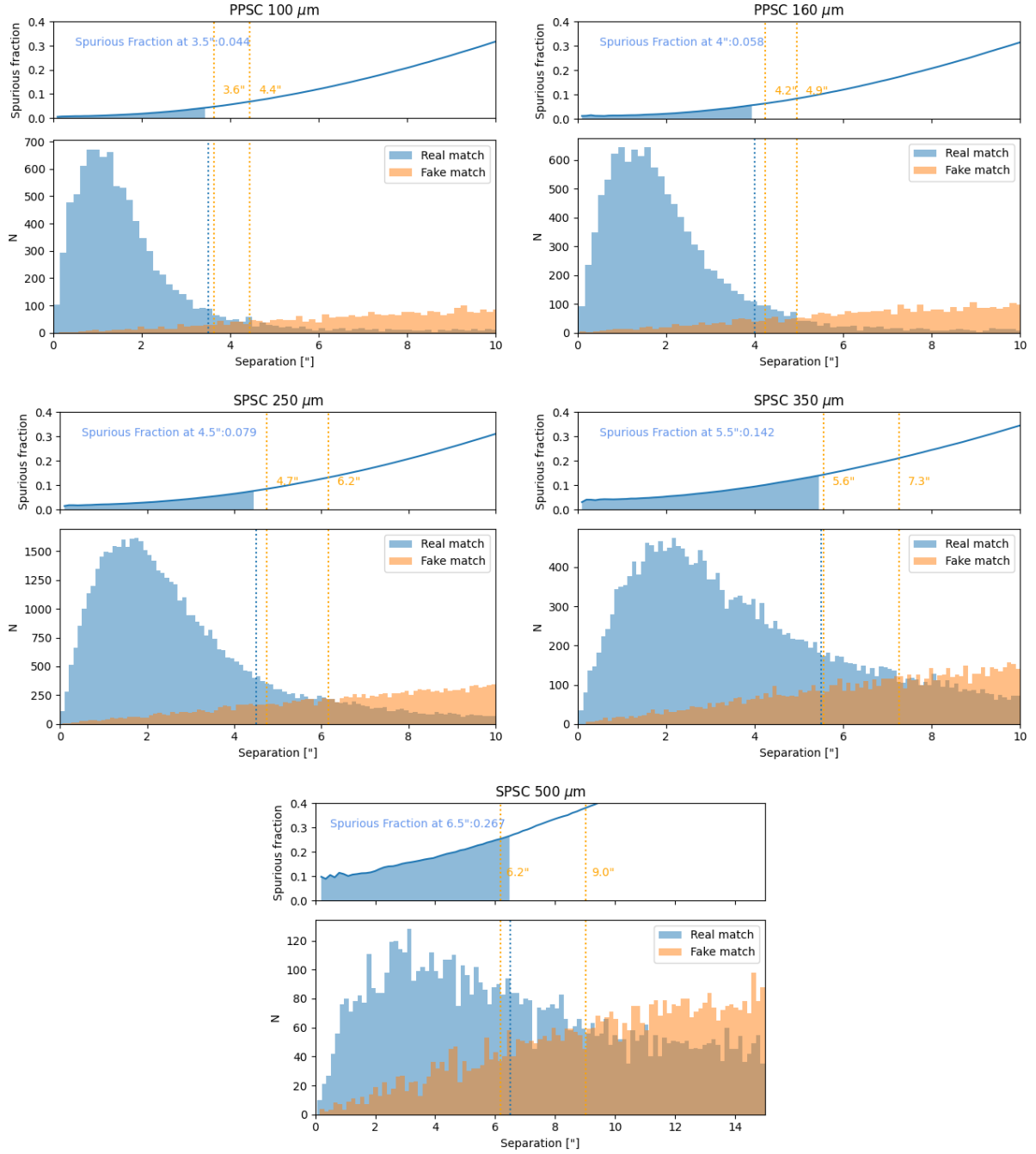


Figure A1. Estimated contamination fraction from spurious matches as a function of matching radii for the PPSC and SPSC data. Blue histograms in each panel represent the distribution of optical – IR angular separation between SDSS DR12 positions in our parent catalogue and the Herschel point source catalogue positions of a given band, while the orange ones are the number of matches as a function of separation between the optical and one of the mock IR catalogue positions. The estimated contamination fractions, calculated as the ratio of the real matches and the average separation distribution of all 15 mock matches, is shown in the smaller panels above. Vertical dashed orange lines mark the radial slices in which the real catalogue is expected to consist of 50% and 100% fake sources, while blue dashed lines represent the search radius we adopted for our catalogue. These radii alongside the estimated contamination fractions are listed in Table A1.

APPENDIX B: CONSISTENCY OF FLUX MEASUREMENTS FROM DIFFERENT CATALOGUES

In order to improve on the sensitivity and photometric coverage of previous studies, and maximize the IRRC parameter

space probed, we have drawn observations from a variety of catalogues, as described in Sect 2. The main drawback of this approach is that the resulting heterogeneous dataset involves sources detected and characterised via differing methodologies. In particular, the H-ATLAS and PPSC/SPSC cata-

Table A1. Matching radii used in each band to incorporate PPSC and SPSC sources into our joint catalogue, and the resulting contamination fractions estimated using Fig. A1.

band	match radius [arcsec]	contamination [%]
PPSC100	3.5	4
PPSC160	4	6
SPSC250	4.5	8
SPSC350	5.5	14
SPSC500	6.5	27

logues were produced using different flux extraction techniques. Thus we compared these data in regions of overlapping coverage and applied corrections when necessary, as described in this Appendix. We note that these adjustments translated to only minor changes in the IRRC statistics, which remained qualitatively consistent with the results obtained using the original catalogue fluxes.

The H-ATLAS survey identified extended sources and extracted their fluxes using apertures of appropriate sizes (Rigby et al. 2011). On the other hand, the PACS and SPIRE Point Source catalogues assumed, as their name indicates, appearance similar to the instrument point spread function for all sources. Indeed, a comparison of H-ATLAS and PACS Point Source fluxes²⁰ using our depth-matched SFG sample, as shown in Fig B1, reveals some 10-20 % average excess flux in H-ATLAS measurements at PACS wavelengths. Below $z < 0.2$, many galaxies are expected to be resolved by the ~ 10 arcsec Herschel PACS beam, suggesting that in our set of galaxies flux extraction with the point source assumption misses some emission at 100 and 160 μm wavelengths²¹. On average, the lower resolution SPIRE measurements appear consistent between H-ATLAS and the SPSC within 1 and 2 σ below and above 400 μm , respectively.

Due to their treatment of extended emission, we considered H-ATLAS measurements the gold standard for our Herschel data, and as the simplest approach to mitigate the differences between the two catalogues, we scaled up our PACS Point Source Catalog measurements to match the average flux levels of H-ATLAS. To calculate the scaling values, we fitted the band-by-band median log-space offsets μ as a function of central wavelength for all five Herschel bands with a linear model:

$$\mu = 0.16 \cdot \log(\lambda/\mu\text{m}) + 0.39. \quad (\text{B1})$$

Substituting 100 and 160 μm into this equation yields a correction factor (in linear space) of 1.19 and 1.10 for fluxes measured in the corresponding band. These were applied to all PPSC fluxes in our catalogue. Since the measured flux offsets for SPIRE data were consistent with 0 within 1-2 σ , we

²⁰ From the SPSC catalogues, we use the TIMELINEFITTER (TML) fluxes, which are the most accurate for point sources. PPSC fluxes were measured using the ANNULARSKYAPERTUREPHOTOMETRY task and apertures out to a radius of 18 and 22 arcsec for the 100 μm and 160 μm bands, respectively.

²¹ On the other hand, we have removed the likely most inaccurate SPSC measurements by omitting sources with the e.g. the blend-flag and large galaxy flag from Jarrett et al. (2003).

did not adjust SPSC measurements. However, to retain information on the uncertainties of the PSC correction/scaling factors (both for PACS and SPIRE) in our SED modelling, we added the error on the median in each band in quadrature to the tabulated PSC flux uncertainties. These were 50, 80, 60, 50, and 30 mJy in the 100, 160, 250, 350, 500 μm bands, respectively. These are a factor of 2-4 larger than the typical 1 σ flux uncertainties of PSC measurements (see Table 1), and they thus dominate the error budget of PSC data used for SED modelling.

A more detailed inspection of Fig. B1 shows that despite the generally weak dependence of flux density ratios on the intrinsic flux density (traced by H-ATLAS measurements), in the highest H-ATLAS flux regime there is an upturn of flux density ratios in almost all bands, possibly related to aperture effects. This suggests that a flux density dependent correction factor for the brightest sources could provide a more accurate correction for PPSC and HPSC flux densities than a simple scalar multiplication. To assess the impact of the inconsistency between H-ATLAS and PSC fluxes, as well as the effect of our scalar correction, we obtained L_{TIR} estimates via SED fitting using SPSC and PPSC fluxes with and without our scaling applied and using only H-ATLAS photometry. We found that all three are consistent with one another within 10%, which is well within the typical 20-30 % uncertainties of our SED models. Ultimately, a homogeneous flux measurement approach would be preferable, but according to our assessment above the impact of the combining fluxes from different catalogues constructed with different methods in practice is not large. Therefore, considering that the impact of the combining fluxes from different catalogues constructed with different methods is in practice not large, and the fact that 65 – 80 % of all PSC measurements lie in the constant flux density ratio regime (as evidenced by the histograms in PPSC and HPSC sources in Fig. B1), for the sake of simplicity we decided to apply the correction to PPSC flux densities in the form of a single value, as described above.

APPENDIX C: PUBLIC DATA RELEASE

In order to aid future studies on the dependence of the IRRC of low- z galaxies on various galaxy properties, we make our data publicly available. Tabs. C1 and C2 show a small section of our data release to illustrate its content, while the online supplementary material, as well as the Zenodo data repository (at <https://zenodo.org/badge/latestdoi/344211798>), contain our entire combined sample of the jointly IR- and radio-detected galaxies. Its content is as follows:

Column (1): SDSS ObjID – Galaxy IDs from SDSS DR12 (corresponding to the SDSS spectroscopic ID when available, otherwise it is the SDSS photometric ID; see SpecFlag in column 5).

Column (2): R.A. [J2000] – Right ascension in degrees from SDSS DR12.

Column (3): Dec. [J2000] – Declination in degrees from SDSS DR12.

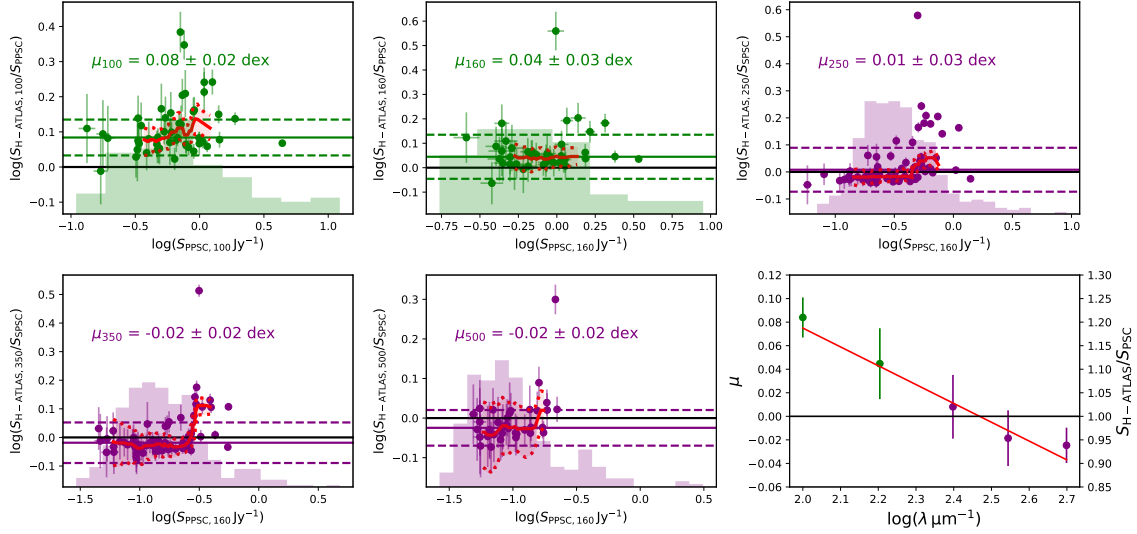


Figure B1. Band-by-band flux density ratios of measurements from the PACS (green) and SPIRE (purple) Point Source Catalogs and photometry from the H-ATLAS survey as a function of H-ATLAS flux densities using star forming galaxies in our combined sample. The median logarithmic H-ATLAS to PSC flux density ratio of any given photometric band is denoted by μ in each panel, and it is represented by a green (purple) horizontal line for PACS (SPIRE) panels. Horizontal dashed lines of the same colour show the 3σ uncertainties of the median value. The 1:1 relation, i.e. a logarithmic flux density ratio of 0, is shown as a black line. Red curves are the running median values. Dashed red lines indicate the 3σ range around the running median. The final panel shows the median logarithmic flux ratio values μ as a function of wavelength, alongside our best-fit relation, Eq. B1, used to derive flux corrections in the 100 and 160 μm bands.

Column (4): z – spectroscopic or photometric redshift from SDSS DR12, see column (5) detected. Column (5): Specflag – Set to 1 if a source was spectroscopically detected, otherwise 0.

Column (6): BPTflag – Flag values of 0, 1 and 2 denote SFG, composite and AGN sources, respectively, as described in Sect. 3.2.1. A value of -1 indicates the lack of high quality spectral line detection, or have no $H\alpha$ detections at the $\text{SNR} > 3$ level in the MPA/JHU value-added catalogue, and consequently in SDSS DR 8.

We note that our redshifts were drawn from SDSS DR 12, therefore a flag of 0 in column (5) in a handful of cases does not correspond to a flag of -1 in this column and vice versa.

Column (7): MIRflag – Set to 1 if a source was labelled as an AGN based on its MIR colours (for details see Sect. 3.2.2), otherwise 0.

We caution users against including sources with MIR flag =1 in their analyses, since our SED templates assumed pure SF across the entire IR wavelength regime. As a result, their $\log(L_{\text{TIR}})$ (and by extension, q_{TIR}) values are likely overestimated. We note that galaxies labelled as AGN based on optical emission lines but not on their MIR colours are still considered to have robust $\log(L_{\text{TIR}})$ estimates.

Column (8): FIRSTflag – Set to 1 for galaxies in covered by the FIRST survey, otherwise 0. It is required to reproduce our depth-matched sample (Sect. 3.3).

Column (9): SEDflag – Set to 1 for galaxies with robust IR SED models, otherwise 0.

Column (10 - 37): S_{band} and E_{band} – flux densities and

their errors as tabulated in the various archival catalogues used to assemble our data (see Sect 2) in Jy. These bands are 22 μm WISE, 60 μm IRAS, 100 μm IRAS, 100 μm H-ATLAS, 100 μm PPSC, 160 μm H-ATLAS, 160 μm PPSC, 250 μm H-ATLAS, 250 μm SPSC, 350 μm H-ATLAS, 350 μm SPSC, 500 μm H-ATLAS, 500 μm SPSC, 1.4GHz flux density (either NVSS or FIRST, see Sect. 2.2), respectively. We note that the IR flux values published here are *not* re-scaled/corrected²²

Column (38): $\log L_{1.4}$ – 1.4 GHz radio luminosity in units of $\log(\text{WHz}^{-1})$.

Column (39): $\log L_{1.4}$ – 1.4 GHz radio luminosity uncertainty.

Column (40): $\log L_{\text{TIR}}$ – Total logarithmic IR luminosity from IR SED fitting (see Sect. 3.1.1) in units of $\log(L_{\odot})$.

Column (41): $\log L_{\text{FIR}}$ – Logarithmic far-IR luminosity from IR SED fitting (see Sect. 3.1.1) in units of $\log(L_{\odot})$.

Column (42): $\log L_{\text{LIR_upp}}$ – Upper $\log(L_{\text{TIR}})$ luminosity uncertainty corresponding to the difference of the 84th percentile and the median of the marginalized $\log(L_{\text{TIR}})$ posterior distributions from our MCMC fits. We note that the errors on $\log(L_{\text{FIR}})$ and $\log(L_{\text{TIR}})$ are considered to be equal.

²² Before SED fitting we carried out the following modifications: (i) we multiplied our 100 and 160 μm PPSC data by 1.19 and 1.10, respectively; (ii) we increased the uncertainty of all PPSC and HPSC measurements by adding 0.05, 0.08, 0.06, 0.05 and 0.03 Jy in quadrature to the tabulated flux density uncertainties of 100, 160, 250, 350 and 500 μm data.

Column (43): $\log(L_{\text{LIR_low}}$ – Lower $\log(L_{\text{TIR}})$ luminosity uncertainty corresponding to the difference of the median and the 16th percentile of the marginalized $\log(L_{\text{TIR}})$ posterior distributions from our MCMC fits. We note that the errors on $\log(L_{\text{FIR}})$ and $\log(L_{\text{TIR}})$ are considered to be equal.

Column (44): $q_{\text{TIR}} - q_{\text{TIR}}$ calculated with Columns (38) and (40).

Column (45): $q_{\text{FIR}} - q_{\text{FIR}}$ calculated with Columns (38) and (41). By definition q_{FIR} is lower than q_{TIR} by $\log(L_{\text{TIR}}) - \log(L_{\text{FIR}})$.

Column (46): $dq_{\text{TIR_upp}}$ – Upper error on q_{TIR} , calculated by propagating the uncertainties in Columns (39) and (42)).

Column (47): $dq_{\text{TIR_low}}$ – Lower error on q_{TIR} , calculated by propagating the uncertainties in Columns (39) and (43).

To select our depth-matched SFGs with reliable SED models one has to select sources with $\text{FIRSTflag} = 1$, $\text{BPTflag} = 0$, $\text{MIRflag} = 0$ and $\text{SEDflag} = 1$ and $S22_{\text{WISE}} > 0.0195$ Jy.

This paper has been typeset from a $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ file prepared by the author.

Table C1. First 13 columns of our publicly available low- z IRRC catalogue. The full catalogue, containing all 9,645 sources in our combined sample, is available as online supplementary material. For a description of all columns, see Appendix C.

ObjID	RA	Dec	z	Specflag	BPTflag	MIRflag	FIRSTflag	SEDflag	S22 _{unWISE}	E22 _{unWISE}	S60 _{IRAS}	E60 _{IRAS}	...
-	[deg]	[deg]	-	-	-	-	-	-	[Jy]	[Jy]	[Jy]	[Jy]	...
1237653651309002970	25.76	13.65	0.0029	1	2	0	0	0	0.1931	0.0006	69.05	0.05	...
1237648720159965190	186.73	-0.88	0.0071	1	2	1	1	0	12.744	0.002	40.68	0.041	...
1237657628439543935	135.11	39.06	0.0582	1	1	1	1	0	1.7040	0.0007	7.43	0.04	...
1237661815485104153	189.14	11.24	0.0074	1	0	0	1	0	0.2148	0.0008			...
1237664853107933399	217.02	32.89	0.1829	0		0	1	1	0.0045	0.0005			...
1237662225149722677	209.64	37.45	0.0115	0	0	0	1	1	1.0615	0.0009			...
1237664336632021061	149.17	30.75	0.0526	1	0	0	1	1	0.0276	0.0009	0.32	0.04	...
1237666245748064333	314.35	17.13	0.0883	0		1	0	1	1.178	0.001			...

Table C2. The last 10 columns of our publicly available low- z IRRC catalogue. The full catalogue, containing all 9,645 sources in our combined sample, is available as online supplementary material. For a description of all columns, see Appendix C.

...	$\log(L_{1.4})$	$\log(\Delta L_{1.4})$	$\log(L_{\text{TIR}})$	$\log(L_{\text{FIR}})$	$\log(\Delta L_{\text{TIR,upp}})$	$\log(\Delta L_{\text{TIR,low}})$	$\log(q_{\text{TIR}})$	$\log(q_{\text{FIR}})$	$\log(\Delta q_{\text{TIR,upp}})$	$\log(\Delta L_{\text{TIR,low}})$
...	[$\log(\text{WHz}^{-1})$]	[$\log(\text{WHz}^{-1})$]	[$\log(L_{\odot})$]	[$\log(L_{\odot})$]	[$\log(L_{\odot})$]	[$\log(L_{\odot})$]	-	-	-	-
...	21.83	0.02	10.5106	10.2103	0.0003	0.0003	2.69	2.39	0.02	0.02
...	21.66	0.01	11.19857	10.91486	0.00001	0.00001	3.55	3.27	0.01	0.01
...	22.54	0.05	12.1011	11.8174	0.0001	0.0001	3.57	3.29	0.04	0.04
...	22.184	0.02	10.7	10.5	0.7	0.3	2.6	2.3	0.3	0.2
...	22.89	0.06	11.08	10.76	0.09	0.09	2.20	1.88	0.07	0.07
...	22.01	0.01	10.7724	10.5129	0.0002	0.0002	2.77	2.51	0.01	0.01
...	22.19	0.08	10.80	10.49	0.09	0.11	2.63	2.31	0.09	0.09
...	23.91	0.02	12.567	12.307	0.001	0.001	2.66	2.40	0.02	0.02